

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 15, 2002, 22:21:07 ; Search time 1836.85 Seconds

(without alignments)
5707.703 Million cell updates/sec

Title:
Perfect score:

US-09-484-577a-3
501

Sequence:

1 actctccagcctctcaccga.....cagaatgcgcatgatcat 501

Scoring table:

IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched:

1797656 seqs, 10463268293 residues

al number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Genemdb1:*
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_ov:*
22: em_or:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_higo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
------------	-------	-------	--------	----	----	-------------

1	92.4	18.4	3526	1	AF273216	AF273216 Rhizobium
2	88.8	17.7	28804	1	SS051197	US1197 Spingomonas
3	88.8	17.7	28804	6	AR068625	AR068625 Sequence
4	86.8	17.3	15552	1	AE004049	AE004049 Xylella f
5	86	17.2	1428	1	NME391263	AF121772 Neisseria
6	86	17.2	4254	1	AF121772	AF121772 Neisseria
7	86	17.2	7824	1	NME391260	AJ391260 Neisseria
8	86	17.2	11381	1	AE002524	AE002524 Neisseria
9	86	17.2	326301	1	NMA622491	AL162757 Neisseria
10	86	17.2	349980	6	AX044033	AX044033 Sequence
11	86	17.2	349980	6	AX044034	AX044034 Sequence
12	80.4	16.0	189050	1	AL646077	AL646077 Ralstonia
13	69.8	13.9	334050	1	AJ414151	AJ414151 Yersinia
14	62.6	12.5	6645	1	BPCYADE	X114151 Yersinia
15	56.8	11.3	7493	1	AF193064	AF193064 Caulobact
16	56.8	11.3	13759	1	AE005779	AE005779 Caulobact
17	56.8	11.3	19883	1	AF062345	AF062345 Caulobact
18	54.2	10.8	15441	1	SC011	AF096823 Streptomy
19	53.2	10.6	208050	1	AL646083	AL646083 Ralstonia
20	51.6	10.3	5120	1	APCLYIAD	X61112 A. pleuropne
21	51.6	10.3	8292	1	AR091872	AR091872 Sequence
22	48.6	10.3	8370	6	AR091872	AP003007 Mesorhizo
23	48.6	9.7	348411	1	AP003007	ME5808 Actinobacti
24	48.4	9.7	4247	1	ACNAPAP	ME5808 Actinobacti
25	48	9.6	7413	1	PAPAPAP	X64558 P. aeruginos
26	47.8	9.6	11588	1	AE004554	AE004554 Pseudomon
27	47.8	9.5	6400	6	E30060	E30060 ABC Transpo
28	47.8	9.5	11273	1	AB023289	AB023289 Pseudomon
29	47.6	9.5	11383	1	AE004761	AE004761 Pseudomon
30	47.2	9.4	299350	1	SMES91786	AL591786 Sinorhizo
31	47	9.4	18857	1	AE006947	AE006947 Mycobacte
32	47	9.4	35420	1	MTCY22610	284724 Mycobacte
33	47	9.4	42741	1	MSGY423	AD000014 Mycobacte
34	46.6	9.3	333500	1	AP003590	AP003590 Nostoc sp
35	46.4	9.3	10029	1	AE009262	AE009262 Agrobacte
36	46.4	9.3	10235	1	AE008350	AE008350 Agrobacte
37	45.6	9.1	10072	1	AE005876	AE005876 Caulobact
38	45.6	9.1	340857	1	AP003010	AP003010 Mesorhizo
39	45.2	9.0	215050	1	AL646057	AL646057 Ralstonia
40	45.2	9.0	264050	1	AL627279	AL627279 Salmonell
41	45.2	9.0	274050	1	AL627279	AL627279 Salmonell
42	44.8	8.9	22991	1	SMES91792	AL591792 Sinorhizo
43	44.4	8.9	3368	1	SCD19	AL592149 Streptomy
44	44.4	8.9	3368	1	S67013	S67013 prtd-Protea
45	44.4	8.9	4561	1	SMASADE	X81195 S. marcescen

ALIGNMENTS

RESULT 1	AF273216	3526 bp	DNA	Linear	BCT 17-JUL-2001
LOCUS	Rhizobium leguminosarum bv. viciae bacteriocin (trza) gene, partial cds; ABC transporter RzcB (rzcB) gene, complete cds; and transporter component RzcD (rzcD) gene, partial cds.				
DEFINITION	AF273216				
ACCESSION	AF273216.1	GI:10952793			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM	Rhizobium leguminosarum bv. viciae.				
	Rhizobium leguminosarum bv. viciae				
	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;				
REFERENCE	1 (bases 1 to 3526)				
AUTHORS	Venter, A.P., Twelker, S., Oresnik, I.J. and Hynes, M.F.				
TITLE	Analysis of the genetic region encoding a novel Rhizobiocin from Rhizobium leguminosarum bv. viciae strain 306				
JOURNAL	Rhizobium journal of microbiology. 47 (6), 495-502 (2001)				
MEDLINE	21360352				
PUBMED	11467725				
REFERENCE	2 (bases 1 to 3526)				
AUTHORS	Venter, A.P., Twelker, S. and Hynes, M.F.				
TITLE	Direct Submission				

GSNTRTPPEWVWVNNMLGCVSLVNLVIGFEPFRARHVLNVCWLIQINDSVILHLTQ
VGTIRVANNVGNPLVWGLVAMTIPRAHILMVNSRSVALTLTLPDLAMVLSYWS
GARLSMVWVVAHATYLLLMGRLRMNIREKGLVGIISTMMPPAFVAVLVMVPAV
HNRVYAGQASNDGRHEQIRMGLPKMLAMPVEVGFPOSALVAVGMSRSGSELSIDS
GELISAVDYGVGLFAFYGTWVLAGLIRAGLIDPDSOGYPLHLAVATVMVNLNRL
VLSQTDNDQVFFMGLTFFALLYRSKQARCA"
/gene="spss"
complement(1955. .3313)
/gene="spss"
complement(1955. .3313)
/function="sphingan polysaccharide synthesis"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="unknown"
/protein_id="AAC44056.1"
/db_xref="GI:1314563"
/translation="MTSSGSLPPGPGCHKAOTGSFVGQAFNVTRSKLHRLGCOMVVG
FGYKRGAVTSFAFTVLVAFSGPAGAGFALTATQMLVILSIGLDTLVRTAV
CVSONRTGAVRTAVRATRAVAGGGLVAMVILFHQOIALNLLGSPAMPDLILMGF
LIPVSAICRLASTTLRGMGDIKQSDWVDPGLGTFELGAGMLAGALLIGVANNPLPSVL
YLVGMVLTMRATIIYRMTDRMAPAPDLRPMLRGFWLITNNNEFYDMRTVYL
AATHGPAEGLFVGVQIASLKLSTSTLHPVPAASYEGQVAKRIARITITL
ALVVAAPLALAVLVAHPHIMGLFGQFTAGAAAMOVYLVGVFVFASSGGVLVWA
IT"
/gene="spss"
complement(313. .5325)
/gene="spss"
complement(313. .5325)
/function="sphingan polysaccharide synthesis"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAC44057.1"
/db_xref="GI:1314564"
/translation="MPDIIVRNOSELDIAIKAKGGERIKLAAGTYSTVTYTTFTT
FVTTISLASKRVNVTTLQIINSSNLVKNVTAAPNPOADXYNANRIGSSNVID
GVTLISGTDVYKSMGMLLRESKGVTKNSVDHPALGNVLTVDGTVLGNNEHD
NRDHTNFAEASNVYIDNNMFTNVPNGEHPDAIQFTNGTKNTNITNNVMO
AIDWIRLDVGVNNESSIPYENITTKNNLVYQNGFTGHVNVGKNVYDSMTIISKD
KAOLANLIVSVGYOPPAQSAFAAQAQVAAQSTGANLLDLPATGIVDQTRV
STDETTALDLTASINMEFVKTGTGELTSDNSROLFALPAFTLNPDKRATAGAY
GQIMGVNOSWGISLRADGELFTVKNAAQSYTATSGAKMTDTRATHTALITTSQK
KAIIVGVNVSSTIIVSTRAYAGLYISGPNVAVSSGVGEIEMKDALSSQIL
ALNAAATASQPTTVKPSLSSILASAVTITGTPVISTALPAATGTTLTGGSASTA
5526. .6473
/gene="spso"
5526. .6473
/gene="spso"
/function="sphingan polysaccharide synthesis"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="glycosyl transferase"
/protein_id="AAC44058.1"
/db_xref="GI:1314565"
/translation="MEASPTDVSLVVAHSAFPIGOCIRGTAAGTAAGTAHLLID
NCGGDEAVVRAEFPHRIYVSEGNIGFEGNRCARHARAPULLLVNPPALPRAPI
DLVAFKAPADAAAGRSYFPNGOLDANETPLTVRDFVVISSTSSPMRGGPAI
DATAEPVNLNGEMVADVARETIDGDEGFTLSEEDLQIRIRAGYSVLVPA
VGVHDTGGHSLSPTRVLTITGRHRIARKEHGHGAVVGTGMAWMAKAYVYIGGL
LGRISPRARANALPDASIVFGQRRWGHWRDVRH"
/gene="spst"
complement(6558. .7460)
/gene="spst"
complement(6558. .7460)
/function="sphingan polysaccharide synthesis"
/codon_start=1
/transl_table=11
/product="unknown"

gene
CDS
/protein_id="AAC44059.1"
/db_xref="GI:1314566"
/translation="MKKILYLVAAVAAALVAGCGSGKEKLDKGVASVDEITIF
ELNAEVAOVPOFGVDRLABQMLORIETIKTAIEQKLDLDPFLIDQKRADE
LITLLLRKIRIGIAOPDPAIDIAOTEAHNPORNTISLEQVFPFGSAAKE
FAPKTLIOIAKLADGVORFAPPTALPELPAKIALPAPQELFIPQOGV
TANVTGTTIOPDLAGDQARELNLRLRERSKADAOLNRLKARSVKQPEYSP
PQOLGGAAAPAAPAPNAQ"
/gene="spsk"
7588. .8646
/gene="spsk"
7588. .8646
/function="sphingan polysaccharide synthesis"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="glycosyl transferase"
/protein_id="AAC44060.1"
/db_xref="GI:1314567"
/translation="MGDMAEATVTEKAGKPLKMLAASGGHLROILDESVMKEH
DYFVTEDTALGRSLAKRHSVALVDHIALQAKIGHPLRMIGAMRLRQSLTIKHK
KPVVITSGAVYFTLAKLSGAKFVHIESFAREDHPAGMKVGIATVYOSA
ALQOTWDALEFDEPRLDPRPKQALPEATVGAATLPPRLVOAVTDLKRAAGLCK
LVQYGDQDLADPGIPDVEIRTIPEFDLQLLLRDADMVCHGCTGSLVATLACGV
VAFPRRDLEHYDDHOEEIAGTFADRLGLHVRDERELGAAYEAAKATEPQLATYDH
TALAGRLRELIQAOMSAKR"
/gene="spsl"
8643. .9509
/gene="spsl"
8643. .9509
/function="sphingan polysaccharide synthesis"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="glycosyl transferase"
/protein_id="AAC44061.1"
/db_xref="GI:1314568"
/translation="MSAPRISVYIPHYNDPDSLRQCLDALQHTIGREAFETIIGDN
SFCGLAAVAAGARAVYTLIEKGAAPARNGAAAEAGEIYAFDSCVYVPGMLAG
GVAHAAPREVGVMHYLPEGRITGAELVLAALAEAGEIYAFDSCVYVPGMLAG
DFERVGEFTVSEDELMCHRAITAGLIDVAPESVCHPPRPDATTLLVTRIORE
LLEFNIERPGRLRLARSTIQLPILPDTAKILTPTGRSRIAAVGTVLRLPWA
GAGLQLLGRPI"
complement(9552. .10940)
/gene="spst"
complement(9552. .10940)
/function="secretion of sphingan polysaccharide"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="unknown"
/protein_id="AAC44062.1"
/db_xref="GI:1314569"
/translation="MAVGSALFLMPGRDVPREDGYFPLAATVPHRAHGRGP
DPTFRASALDRPLDRRERRETRAPPLATPTPTQVADRSSFAGLEVALRI
SAIESQSRVVLVGERGIGKSLIHLVTVARBSSTVYATGANNFADLRAVRI
DVPLFRGVSPNAGESEGSADRIPTGSFSGELADICADITGTRVLIIDEVD
VTPDFRQOVALIKNISDSARQVLTAVSSNLDELVGVYSIRNITGLPMPRIE
ETVVOEMIALGETASGRFDPNLTIRIHLALGSPYFARLCHHALEALDQTCGLAV
IGHRLRDLADITIEIGRMSPRVLEMRKVGGRYPALVALGASRSRDTGQGV
VELLPSTTTAAVVEQELDALISROLGIESDAQOGERFRFNRNDSLIVYILMGRILTI
DSQTLLEALAAV"
/gene="spsf"
11569. .12867
/gene="spsf"
11569. .12867
/function="sphingan polysaccharide synthesis"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="unknown"
/protein_id="AAC44063.1"


```

sequence"
/codon_start=1
/transl_table=11
/product="putative outer membrane protein"
/db_xref="GI:6900416"
/transl_table="MTLNLIMQDYGISVCLTLPYLOHELFESANKSYFSKYLIPVS
LEFTPLSPSVASFTLPEAMRAAOHSDFOASHYQDRAVRAOQAAALPLHVS
NASTYQOPSPISSTRETQKMSVQVQGLDPAKFAQYROSREPTQAAEQDFAAREL
LKAASSTYFNVLSPRTVAHAHAKRAYAOQAOALFNKGAATLALDIEHAKAGYN
ALANHEYRMQOALOSSGALRAQNSRTYTAHAYGONMLTSSNONNDYRKG
LSNNHEYRMQOALOSSGALRAQNSRTYTAHAYGONMLTSSNONNDYRKG
MSYVOLNPLTYGELSGKHEAEAYGAQOLATERHITLARAQATBESGAR
OIAOEVRLESSLRLKSTETGQOYGIIRNLEVIKAOEYVAOEKLAQARFMALAY
LRLVKSGLGLEVFPAE"
complement(5707..6042)
/gene="rse1"
/complement(5707..6042)
/gene="rse1"
/complement(5707..6042)
/note="homology to putative transposase of CAB44496;
designated NM1992 in the Sanger Center 22491 genome
sequence"
/codon_start=1
/transl_table=11
/product="putative transposase"
/protein_id="CAB72028.1"
/db_xref="GI:6900417"
/transl_table="MSAOVGNRLIAPMYONTMTGVFEAFPOCCLPALTQKSVILL
DPAHFHMGVLRMAEKLYKVLPLAPLPLNIEKWMANIKRYLRLTVLSDVAFED
ALLSTFDPEN"
6514..6619
/note="c106, Correia element"
complement(6787..7824)
/gene="relA"
/complement(6787..7824)
/gene="relA"
/function="synthesis of ppGpp, a mediator of the stringent
response"
/note="homology to GMP-pyrophosphokinase of P1185;
designated relA (NM1991) in the Sanger Center 22491 genome
sequence"
/codon_start=1
/transl_table=11
/product="GMP-pyrophosphokinase"
/protein_id="CAB72029.1"
/db_xref="GI:6900418"
/transl_table="MAESKEDLAAAFKTELFPNDIIVLTPHGKVLSTPTGATPIDFA
VALHSIGRCRGAVEGQIVLPSTLENGQVETITAKEGHSVNMILEGKVNKSKA
IGIRRAYIROQADNIVREGVQLDKQLKLPKPLQELANLGGKRPEDLXTYVQ
GEISNRALQKAGCTLNPPVSETTIVKQSKIRKGGNGVLIQEDSLMTTLAC
KRAPDDIVGFTVRDGI SVHRKTCPSFRLHAEKVLDSMAALQDGVFAVDIE
IKAOBSGLRVDALNRHKNLNTAVQOSDLKESMFTLEVQVDTLPRVLASLG
DVKGVLSVTRL"
BASE COUNT      1853 a      1854 c      2065 g      2052 t
ORIGIN
Query Match      17.2%; Score 86; DB 1; Length 7824;
Best Local Similarity 54.8%; Pred. No. 1.4e-07;
Matches 170; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
QY      62 acccggaagagctgctcctcctcgcgccgacctgagaatgtcgagaagcgccatct 121
      || || || || || || || || || || || || || || || || || || || ||
Db      2606 ACGGCGGAAGAACAGGCTTTTCCCGCATTTGGAACTGACCATACCCGCTCT 2665
QY      122 ccaacgcgagactcaagcgccctgtctgtctgtctactactcgccgtgagctg 181
      || || || || || || || || || || || || || || || || || || || ||
Db      2666 GCCCCTCCGAATAGGCGGCGCTTTTATATGCGCTTTTGGCTTTGTGTGG 2725
QY      182 gcgggctcgcagatcgatctgtctgtcatccagaaagatcgctgcggggac 241
      || || || || || || || || || || || || || || || || || || || ||
Db      2726 TCCGCTTCGGGAAATGATATGTGTGCGCGCTTCGCAAAACGCTGTCGGCGG 2785

```

```

QY      242 cgtgtaagctgtgtcaagccgctcgagctgagctgtgctgggcccactatgctcgat 301
      || || || || || || || || || || || || || || || || || || || ||
Db      2786 CGGAGCAAAACCATCCAGCCGCTGGAACAGGTGTGTATGAAGCGGATGTGCGGAC 2845
QY      302 ggcacaacgctcaagcgcgagatctgtatcgagctgatatccatcgcggtgtgtg 361
      || || || || || || || || || || || || || || || || || || || ||
Db      2846 GCGCAGCATGTGGAACAGGAGAAACGCTGCGGAACTGGAGGCTGTGGAAACAGACG 2905
QY      362 gatgtgcgc 371
      || || || || || || || || || || || || || || || || || || || ||
Db      2906 GATGTGTGTG 2915

RESULT      8
AE002524/c 11381 bp DNA linear BCT 25-MAY-2000
LOCUS      Neisseria meningitidis serogroup B strain MC58 section 166 of 206
DEFINITION the complete genome.
ACCESSION      AE002524 AE002098
VERSION      AE002524.1 GI:7226991
KEYWORDS
SOURCE      Neisseria meningitidis MC58.
ORGANISM      Neisseria meningitidis MC58.
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
1 (bases 1 to 11381)
Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C.,
Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F.,
Dodson, R.J., Nelson, W.C., Gwinn, M.L., Deboy, R., Peterson, J.D.,
Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D.,
Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E.,
Cittone, H., Clark, E.B., Cotton, M.D., Uterback, T.R., Khouli, H.,
Qin, H., Vamathevan, J., Gill, J., Scarlato, V., Masignani, V.,
Piazza, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R.,
Rappuoli, R. and Venter, J.C.
Complete genome sequence of Neisseria meningitidis serogroup B
strain MC58
Science 287 (5459), 1809-1815 (2000)
107103075
2 (bases 1 to 11381)
Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C.,
Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F.,
Dodson, R.J., Nelson, W.C., Gwinn, M.L., Deboy, R., Peterson, J.D.,
Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D.,
Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E.,
Cittone, H., Clark, E.B., Cotton, M.D., Uterback, T.R., Khouli, H.,
Qin, H., Vamathevan, J., Gill, J., Scarlato, V., Masignani, V.,
Piazza, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R.,
Rappuoli, R. and Venter, J.C.
Direct Submission
Submitted (17-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
1. 11381
Location/Qualifiers
/organism="Neisseria meningitidis MC58"
/strain="MC58"
/db_xref="taxon:122586"
/note="serogroup: B"
165..976
/gene="NMB1736"
/note="This region contains an authentic point mutation,
causing a premature stop, and is not the result of a
sequencing artifact; This region contains an authentic
frame shift and is not the result of a sequencing
artifact; similar to PID:1653908 PID:1651692"
complement(1140..2543)
/gene="NMB1737"
/complement(1140..2543)
/note="similar to PID:1139569 percent identity: 48.26;
identified by sequence similarity: putative"
/codon_start=1

```



```

/gene
CDS
/translation="MTLNLIMMODYISVCLTLPYLQHELFSAKSYSKYILPVS
LFTPLPLSBSVSAFTLPEAMRAAOHSHADFOASHYORDAVRAROOAKAFLEHVA
NASYQORPSSISSTRETQSVYSGOTLFLPAKPAQYROSFEPTQAOAORPDAREL
LKVAESYFVNLISRDYVAHAAEKAYAOOVRQOALFNKGATALDIDHAKAGYEL
ALAOEIVALEKOTYENOLDYDLDSCOTGEAIDTALYARILPKLERYISIDEMORIA
LSNNHEYRMQIALOSSGALRAONSRYPVSAHVYQNNLTYSKQNDYHRYKGS
MSYGVOLNPLTYGELSGKIHAEAYGAEOLITERRIKLAVROAYTESGAARY
QIMAOERVLSSRLKSTEGQOYGRNLREYIRARQEAQAEOKLAOARYKEMLAY
LRLVESGLTEVFAE"
complement(2612..4039)
/gene="NMB1738"
/gene="NMB1738"
complement(2612..4039)
/note="similar to GP:558154 percent identity: 65.47;
identified by sequence similarity; putative"
/codon_start=1
/translation="secretion protein, putative"
/product="secretion protein, putative"
/db_xref="GI:7226993"
/translation="WFFSALKSFLSRYIIVRWVAVRDLKPKKRTAEBOAFPLAHL
ELDTVPVSAAPKWAARIMAFALLALMSWFGKIDIVAASGKTVSGRSKTIOPLET
AVKVAHVRDGOHKOGETLAELAVGSDSDVVOSEALQAOISLRYPVAVLALES
RIVPHIDMAQARSLGSDADVOAIOVAHOYMAAOQOALOSALRGHAELOSQA
OEOKLIVSGAIEEOQKTADYRRLRADNFIHAFLEOOSKYSNMNDELSRGOAROI
AATAOAEONRVLTQNLKRDITLALRANQIDITRGQIDKAKOQOOLITQSPADGT
VOELATYTGCVQAQAQKMYIAPDDKMDVEVLNKGDFVEGQDAVAKTESPY
TRIGYLTGVKVSVDHVSHEQGLVYTAVVSIDKHTLNLIDGKAVNLGAMNTAETK
TGRRLVLDLSLTQTKLDESFRER"
4275..4628
/gene="NMB1739"
4275..4628
/gene="NMB1739"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/translation="hypothetical protein"
/product="hypothetical protein"
/db_xref="GI:7226994"
/translation="MTINDLRHFLGLAIFLSIGYGTMLLMLVFAVKKPKRN
PKLISGLQTAAILMLFVWILPKQFGPIKEIQAEEDRKYEAEAVFNEQCKTAG
ERTTRRTYMKGLCC"
4752..5162
/gene="NMB1740"
4752..5162
/gene="NMB1740"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/translation="hypothetical protein"
/product="hypothetical protein"
/db_xref="GI:7226995"
/translation="MGNTAMMWSMNPITPILFGIOVKPIPLNQIFNHIIPARYATFE
NNVDSKLRRHWAGATIRIIDRODEVIAKTITVFEKGLDGTGARMPWKRAILCNK
ERLTSEPLSDVLVLPKYLIRPLIYLSLRDD"
5233..6575
/gene="NMB1741"
/note="this region contains an authentic frame shift and
is not the result of a sequencing artifact; conserved
hypothetical protein, authentic frameshift; identified by
Glimmer2; putative; conserved hypothetical protein,
authentic frameshift"
6620..6910
/gene="NMB1742"
6620..6910
/gene="NMB1742"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/translation="hypothetical protein"
/product="hypothetical protein"
/db_xref="GI:7226996"
/translation="WNTIQGHIYGVKKINQIEKSFDCCLIVAPPLDSSGNALG
SSTTEYDGSANFEQFRNAQFPEANLNVETVTGKTQKLVIGFLYKKG"
6915..7112
/gene="NMB1743"
6915..7112
/gene="NMB1743"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/translation="hypothetical protein"
/product="hypothetical protein"
/db_xref="GI:7226997"
/translation="WQRYVVOVSSTGDFLYLSPETGDIHTKLITNADYFPEEAI
NAGLEIGNQYEFVFGFLMD"
7120..7404
/gene="NMB1744"
7120..7404
/gene="NMB1744"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/translation="hypothetical protein"
/product="hypothetical protein"
/db_xref="GI:7226998"
/translation="WFGRLKNAPSITAKHLENKKEINTCRKYGAKLAVVTAAPL
ALAAHANATLPDPAKNALAEAKADGEMAGIIVGIPALFVSVIRKVM"
7411..7689
/gene="NMB1745"
7411..7689
/gene="NMB1745"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/translation="hypothetical protein"
/product="hypothetical protein"
/db_xref="GI:7226999"
/translation="MYQVGNKCLEKHOENLYFSLVPRIKENGOIVPREYNGSLMK
MSDQPLRLULECSPEKMDLOSGLGTGWIIFGILASVYFSLKRVLK"
7817..8107
/gene="NMB1746"
7817..8107
/gene="NMB1746"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/translation="hypothetical protein"
/product="hypothetical protein"
/db_xref="GI:7227000"
/translation="MKNKNERDEFFYISNSDLKLSSEYDPRLPYVYCYLKEGTGLKN
FSDMCHNFENINENSCFEIKFDSPFIIGNKIDVSDSNFFSVSFEFC"
8148..9653
/gene="NMB1747"
8148..9653
/gene="NMB1747"
/note="similar to GP:3758871 percent identity: 86.60;
identified by sequence similarity; putative"
/codon_start=1
/translation="hypothetical protein, putative"
/product="hypothetical protein, putative"

```

Query Match 17.28; Score 86; DB 1; Length 11381;
Best Local Similarity 54.88; Pred. No. 1.3e-07;
Matches 170; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

OY	62	accgcggaacagatcgtgcctctctctccgacgcgccttcgaatatctgcgaagccgcgcacat	121
Db	3943	ACGGCGGAAGACACAGGGCGTTTGTGGCCCGCATTTGGAACTGACCGGATACCGCGGTCT	3884
OY	122	cccaacacgagactcaacacgagccgcctctgctctgctctgtctctatctgcgcgttgagcttg	181
Db	3883	GCCCGCTCCGAATAGGGCGCGCGCTTTATTTATATGGCGTTTGGCTTTTGGCTTTGTGGG	3824
OY	182	gcggagctctcgcacgagatcgatcatcgctctgcatccagaagatcgtgcgggcgcac	241
Db	3823	TTCCTGGTTCCGGCAAAATGCAATTGTGTGGCGGGCGCTTGGGGCAAAACGCTGTCGGGGCGG	3764
OY	242	cgtgtaaagctggtctcagcgcgtcagagtcgagtcggtgtgtgcgggcacatcattctcgcgat	301
Db	3763	CGCAGCAAAACCATCTCCACCCCTGTGAAACGGGGGTGGTTAAAGCCGTACATGTGGGCAC	3704
OY	302	ggccaaacgcgtcaaaagccgcgcagagattgatctcgaatcgaatccattctcgcggtgtgttg	361
Db	3703	GGGGCGCATGTGTGAACAACAGGGAACGCTGGCGGAACGTGAGAGCTGTGGGAACACACAGC	3644
OY	362	gattgttcgcg	371
Db	3643	GATGTGTGTC	3634

	RESULT	9
NMA622491/c	LOCUS	
326301 bp	DNA	linear BCT 04-DEC-2000
DEFINITION	Neisseria meningitidis serogroup A strain Z2491 complete genome;	
ACCESSION	segment 6/7.	
VERSION	AL162757.AL157959	
KEYWORDS	ALI62757.2 GI:7380371	
SOURCE	Neisseria meningitidis Z2491.	
ORGANISM	Neisseria meningitidis Z2491 Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.	
REFERENCE	1 (bases 1 to 326301)	
AUTHORS	Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C., Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T., Davies,R.M., Davis,P., Devlin,K., Feltwell,T., Hamlin,N., Holroyd,S., Jagsels,K., Leather,S., Moule,S., Mungall,K., Quail,M.A., Rajandream,M.A., Rutherford,K.M., Simmonds,M., Skellton,J., Whitehead,S., Spratt,B.G. and Barrell,B.G. Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491	
TITLE	Nature 404 (6777), 502-506 (2000)	
JOURNAL	20222556	
ONLINE	2 (bases 1 to 326301)	
REFERENCE	Parkhill,J.	
AUTHORS	Direct Submission	
TITLE	Submitted (30-MAR-2000) Submitted on behalf of the Neisseria	
JOURNAL	sequencing team, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk	
COMMENT	Notes: Details of N. meningitidis sequencing at the Sanger Centre are available on the World Wide Web. (URL, http://www.Sanger.ac.uk/Projects/N_meningitidis/).	
FEATURES	Location/Qualifiers	
source	1..326301	
	/organism="Neisseria meningitidis Z2491"	
	/strain="Z2491"	
	/db_xref="taxon:122587"	
	/note="serogroup: A"	
	196..311	
	/note="REP 2; hmms hit to HMM REP 2 (1 - 109), score:	
	87.43"	
repeat_unit	434..453	
	/note=">= 90% match to ATTCGNNNNNGGAAT"	
repeat_unit	/label=dns3	
	498..560	
repeat_unit	/note="Correa element; hmms hit to HMM Correa (1 - 62)	

```

repeat_unit
score: 78.71"
/label-correia
562. .604
/notes="Correia element; hmms hit to HM Correia (114 -
156), score: 51.92"
/label-correia
complement(667. .676)
/notes="Core DNA uptake sequence: gccgcctgaa"
/label-DUS
complement(668. .1813)
/genes="dapE"
complement(668. .1813)
/genes="dapE"
/EC_number="3.5.1.18"
/notes="NMA1730, dapE, succinyl-diaminopimelate
desuccinylase, len: 381aa; similar to many eg. SW:P24176
(DAP_ECOLI) dapE, succinyl-diaminopimelate desuccinylase
from Escherichia coli (375 aa) fasta scores; E(): 0, 56.9%
identity in 371 aa overlap. Contains Pfam match to entry
PF01546 Peptidase_M20."
/codon_start=1
/transl_table=11
/product="succinyl-diaminopimelate desuccinylase"
/protein_id="CAB84958.1"
/db_xref="GI:7380372"
/db_xref="SPTRMBL:Q9JTL0"
/transl_table="MTEQSLAEKALISRPSPYDPRDCOKLVERLYKIGFAEEL
HGDKRNIMLRGKVPVYCFAGHDVPTGPKWDSPPFEPERDGRYKRGADM
KTSIQFYACARFAVEAHDEHOGSIALITDEGDALDGTWKVDYLRKAGELDYC
IYGEPPAVADKCDMTKNGRSGISGNTLVKKGCHIAVPHLAIPVTFAPALLETLO
EYWDGKNKYPPFSTFOISNNGICGATVPIGELVAFNPFSTFSEAGLKGQVHAI
LKHGELQDLOWSGSGPFLNQAGKTIVDAVAALAEVCGIEALSTGSGDGFKA
IKKEIIEGDSNATHTQINENVRDDIPKLSAVYEGILARLGNAY"
complement(836. .1786)
/genes="dapE"
/notes="Pfam match to entry PF01546 Peptidase_M20, score
252.50, E-value 5.7e-72"
complement(1863. .1872)
/notes="Core DNA uptake sequence: gccgcctgaa"
/label-DUS
complement(1927. .2547)
/genes="NMA1731"
complement(1927. .2547)
/genes="NMA1731"
/notes="NMA1731, conserved hypothetical protein, len:
206aa; similar to SW:P44013 (Y552_HAERIN) hypothetical
protein from Haemophilus influenzae (207 aa) fasta scores:
E(): 0, 59.6% identity in 203 aa overlap."
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAB84959.1"
/db_xref="GI:7380373"
/db_xref="SPTRMBL:Q9JQ23"
/transl_table="MLTPKSCDLFNIPFQFSQLKKYQSPESIPQIKADYKQWQMOQ
LIDQVAELGAFAPAPRIERMCNGQVAFHFAFYKDYQYNSAIIYSILNRRRLSV
SLDMHWKFKADVSPFIALPDYRNWLDNEDKXYSFPMHGAASESDYRTVAQONESDR
KLQNDPDKCNIETRDDLRGQDVAAKMIAEIVEDLLPLYEACGK"
complement(2551. .2558)
complement(2570. .2573)
/notes="Core DNA uptake sequence: gccgcctgaa"
/label-DUS
complement(2581. .3090)
/genes="NMA1732"
complement(2581. .3090)
/genes="NMA1732"
/notes="NMA1732, conserved hypothetical protein, len:
169aa; similar to SW:P43971 (Y243_HAERIN) hypothetical
protein from Haemophilus influenzae (172 aa) fasta scores:
E(): 0, 49.7% identity in 169 aa overlap."
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"

```

```

score: 78.1"
/label=Correia
562..604
/note="Correia element; humf5 hit to HMM Correia (114 -
156), score: 51.92"
/label=Correia
complement(667..676)
/note="Core DNA uptake sequence: gccgcgtcgaa"
/label=DUS
complement(668..1813)
/gene="dapE"
/gene="dapE"
/EC_number="3..5.1.18"
/note="NMA1730, dapE, succinyl-diaminopimelate
desuccinylase, len: 381aa: similar to many eg. SW:P24176
(DAP_ECOL1) dapE, succinyl-diaminopimelate desuccinylase
from Escherichia coli (375 aa) fasta scores; E(): 0, 56.9%
identity in 371 aa overlap. Contains Pfam match to entry
PF01546 Peptidase_M20."
/codon_start=1
/transl_table=11
/product="succinyl-diaminopimelate desuccinylase"
/protein_id="CAB84958.1"
/db_xref="GI:7380372"
/db_xref="SPTREMBL:Q9UTL0"
/translation="MTEDPSGLANALISRPSTVPPDRCCQKLLVERLYKIGFAEEL
HGDGKNTLMIRGTGVYVPCFAGHDVVPVGTGEKWDSEPFETTERDGRLYGGAADM
KTSICAGLAFPAFVAEHDHGSIALILTSDDEGALDGTIKVVDLAKARGLIDYC
ITGEPVADKICGMDIKNGRRGSLGMLTVKQOGHATYPHLAINPHTFAPALLELO
EYWDGNNKTFPTSPSIOISINMGTAQNIIPGLNATFNPFRSTSEGLKRVNHI
LDKHQVQDLOLQSGQPLTPAGKLTIDVARAIATCSIEARLSTETGSGSDRFIKA
IAKELIEGSPNATTHQIENNRFLDDIPKLSAYEGSLARLAGNAV"
complement(836..1786)
/gene="dapE"
/note="Pfam match to entry PF01546 Peptidase_M20, score
252.50, E-value 5.7e-72"
complement(1863..1872)
/note="Core DNA uptake sequence: gccgcgtcgaa"
/label=DUS
complement(1927..2547)
/gene="NMA1731"
complement(1927..2547)
/gene="NMA1731"
/note="NMA1731, conserved hypothetical protein, len:
206aa: similar to SW:P44013 (Y552_HABIN) hypothetical
protein from Haemophilus influenzae (207 aa) fasta scores:
E(): 0, 59.6% identity in 203 aa overlap."
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAB84959.1"
/db_xref="GI:7380373"
/db_xref="SPTREMBL:Q9J0Z3"
/translation="MLTPKSCDLEFNTLPFQFSQLKRYQESIPQIKADYKENNOIQO
LIGQVAALGAPAPAPRIEHCWGVRAHFPAFYKFEQKNSALISILNRRRSV
SLDWHYKADVSALIDPYRYRINDTETKYSFEDMMHGAESXDYRTVAQQNSDR
KLQNDPFRICIKHIERDQGRDVAKWIMLEYEDLLPLVEACHGK"
complement(2551..2558)
complement(2570..2579)
/note="Core DNA uptake sequence: gccgcgtgaa"
/label=DUS
complement(2581..3090)
/gene="NMA1732"
complement(2581..3090)
/gene="NMA1732"
/note="NMA1732, conserved hypothetical protein, len:
169aa: similar to SW:P43971 (Y243_HABIN) hypothetical
protein from Haemophilus influenzae (172 aa) fasta scores
E(): 0, 49.7% identity in 169 aa overlap."
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"

```

/translalion="MLTNIARKIFGSRNDRLLLKQYKRSVARINALEQOMALSDADIC
ATAEFEROLRDLADGOTLIDGLIPEAPFAVCBSARSTIGMHPHIVLGNFLVGLVGTISD
RGEKGTIVATLAVYNLAKGAGCHVYVYNLADSRDGMPEPLINFLVGLVGTISD
MOPERONAYADITTCYGNNEGEPDLYLDMNWTODYDKQRELNFAVDEVDSDILDEE
ARTPLLISQADNDITQLVYIMNTVPVPHLVRQTEGEGEDGVDEKABHGVILSEGHEDIE
ABOILITQGLLAENDSLYSANNTALMHMLALMHLNHTPEFHDHVOYIDGSEIVADEE
TGLMSGRMSGELHQAWEAKGEVIEKRENTOLITITONERLYTTKLSMTGADPE
AFEPQSVNLETVLIPTRNVPQRKDFOITRSAAKEKREAVYKIDDEBCHKRGQVYVVG
TTSINSEVLSHLLOKAGLPHNVLNAKKHEKREBALIYVAGAGVCAITVATNAGKGTID

Query Match	Best Local Similarity	17.2%	Score 86:	DB 1:	Length 326301:
Matches 170:	Conservative	0:	Mismatches 140:	Indels	Gaps
QY 62	accgcggacgagctggcctctcctgcgcgcgcctcgaatgtgcgacgcgcacatc	121			
Db 273549	ACGGCGGAGAAACAGCGCTTTTGGCCCGCGCATTTGGAACTGACCGATACGCGGGTCTCT	273490			
QY 122	ccacgcgagagctcaacggccgcctctgctctgcttcttaactgcgcgctggcgttgg	181			
Db 273489	GCCGCTCCGAAANGCGCGCGCTTTTATTAAGGGCTTTGCGCTTTTGCTTTGTTGGG	273430			
QY 182	gcggggtctgcgagagctacacatcgcttgccttcgcatccaagaagatgctccggcgac	241			
Db 273429	TCCTGGTTGCGAAATGATATTGTGGGGCGGCTTGGCGAAACGCTGTGCGGCGG	273370			
QY 242	gcttgaagctggttcaagccgcctcgaagtcgagtcgtgtgtgcgccaactatgtccgcat	301			
Db 273369	CGCAGCAAAACCATCCAGCCGCTGGAACCGGTGGTTAAGGCGCTGATGTCGCGCAC	273310			
QY 302	ggccaaacctcaagccgcgagatctgtatcgagctcgatccatcctgcgcggtgtgtg	361			
Db 273309	GGGCAGCATGTGTAACAGGAGAAACGCTGGCGGAACTGAGCGCTGTGGAAACAGACAGC	273250			
QY 362	gattgttgcgc	371			
Db 273249	GATGTGGTGC	273240			
RESULT 10	AX044033/C				
LOCUS	AX044033	349980 bp	DNA	linear	PAT 24-NOV-2000
DEFINITION	Sequence 112 from Patent WO006791.				
ACCESSION	AX044033				
VERSION	AX044033.1	GI:11342917			
KEYWORDS					
SOURCE	Neisseria meningitidis.				
ORGANISM	Neisseria meningitidis.				
REFERENCE	1 (bases 1 to 349980)				
AUTHORS	Piazza,M., Hickey,E., Peterson,J., Tettelin,H., Venter,J.C., Masigian,A.V., Galeotti,C., Mora,M., Ratti,G., Scarselli,M., Scariato,V., Rappunoli,R., Frazer,C.M. and Grandi,G.				
TITLE	Neisseria genomic sequences and methods of their use				
JOURNAL	Patent: WO 006791-A 112 09-NOV-2000;				
FEATURES	CHIRON CORPORATION (US) ; THE INSTITUTE FOR GENOMIC RESEARCH (US)				
SOURCE	Location/Qualifiers				
	1..349980				
	/organism="Neisseria meningitidis"				
	/db_xref="taxon:487"				
	/note="sequence too long, cut in 8 pieces, -seq 1: 1 to 349980 349980 bases-seq 108: 300001 to 649980 349980 bases-seq 109: 600001 to 949980 349980 bases-seq 110: 900001 to 1249980 349980 bases-seq 111: 1200001 to 1549980 349980 bases-seq 112: 1500001 to 1849980 349980 bases-seq 113: 1800001 to 2149980 349980 bases-seq 114: 2100001 to 2272325 172325 bases"				
BASE COUNT	87189 a 93501 c 84627 g 84663 t				
ORIGIN					

predicted by Homology
predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="PUTATIVE VGR-RELATED PROTEIN"
/protein_id="CAD17326.1"
/db_xref="GI:17430643"
/translation="MDEKTLQHLFAPARLYALGEGCPILAELAVEMWLGRELSLEF
EWRVVAASANARFALSEPTGCGVTVTTLLADGTQARKTGLIRQAEQLGADGSLARL
IAQFRETREYTRLLAEGLFTTVEDEQAPSGHALIFADSPRLAEDTAAAGGI
RTHRAHSEERDAIQALICHSTRTAVGVAVAAADAKARAFARHPSRRCGIAAGSDA
YLSISPLAPDPAANQRIAEQVMESEARARLFIRGSVTRTSGSTRKIYVCPHLK
VDGPPPLDLVEHGIINNLTADQATLARLTIGSGTTRTGTGRTGRTGRTGRTG
MAPHEVIERETPTADLTAAARAHGVAOPFRAGDARPRPVSHPTDGLTVAEPTAG
VOSATVVGPNGETEASGDAEHHTTPRGQVRRFPWQGEPRDSTGWLVAORAGA
GEMQWMLPRIGOEVIYKQEDDIDQPVVIGALYNGQEGAGIAPPGQVAKVGRGAD
LVFDSDQLRTQATQAYSQLNGLIHQDNKRSFSGQGFELRTDGGVAVGSA
GVLITTRDATSGKAVPTGDNAAGIALIRQAKOLTAISGCGATTHTOTALSTANDSA
PLAOEKASGMVQKALDAQAODAAAGNTGCKVPHQSEATVQIARAGIATAGQ
DLQFANGSEVYLAGSDTNAVGKQARHAGQICGVAAGLSKAGDNDIGLDTAGQDN
LDVOAODALGILSKDLKIVASNLHYDPAKAKRIRLATAGASITLLEGNIIVETPG
RITYKAORSPFAGVYQSYPLPVPQGVCEVLRAMGKHAFANKNG"
1501. 2865
/note="RS50+ or RS06148
PROBABLE LONG DIRECT REPEAT 50 +
predicted by Homology"
/evidence=not_experimental
2999. 3835
/gene="RSP0176"
/note="RS04688"
2999. 3835
/gene="RSP0176"
/function="miscellaneous; hypothetical/global homology"
/note="Product confidence : hypothetical
Gene name confidence : hypothetical
predicted by Codon_usage
predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="CONSERVED HYPOTHETICAL PROTEIN"
/protein_id="CAD17327.1"
/db_xref="GI:17430644"
/translation="MATRRQIOPDRHRTPAWDDYPAADAVEQLSALPPLDYVALVDN
AYDPDAHRLRRFPDMPPRSLYAGRYEGPGLDEIAPTLMRLDPAERTRLDLFLQ
ETSGKPMLSFLHANAPADPLTHLCNOMEAVDHGTAFLRLADTNALAVLOVEPDA
QORFLSDIQWMCFRDGSIRHSHATGGGDAESPYRCTAOMQOLALALAPGAML
NLVONSPHIFGDLTGLPSQAYECIRMAIQAIPGNPTHDAVALRLIASALNEGLIÖH
AA"

gene
4412. 6634
/gene="RSP0177"
/note="RS04689"
4412. 6634
/gene="RSP0177"
/function="miscellaneous; hypothetical/partial homology"
/note="Product confidence : hypothetical
Gene name confidence : hypothetical
predicted by Codon_usage
predicted by Homology
predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="HYPOTHETICAL PROTEIN"
/protein_id="CAD17328.1"
/db_xref="GI:17430645"
/translation="MSYLELGPNPALITVDEQARFLACRMEQEFHCSTEFHLSEFF
DGTNNKRTDTPROASHVVARLFDIFEEQHOIRIYVIGISTPEKETGIDGDRHAR
AGAGAGGGEARIMWALLOVYNALYAYYTQSLSMANGVDELSTVHOTSPLNNGICA
VTAGRDEVEQIANAKNSEMYITAAETVLTFRPHERALLRÖRREYISERTKALING

gene
6745. 7527
/gene="RSP0178"
/note="RS04690"
6745. 7527
/gene="RSP0178"
/function="miscellaneous; unknown"
/note="Product confidence : hypothetical
Gene name confidence : hypothetical
predicted by Codon_usage
predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="HYPOTHETICAL PROTEIN"
/protein_id="CAD17329.1"
/db_xref="GI:17430646"
/translation="MTOYLHITTTAVARIITRTGTLIAHPEALGRVARRHGAFEV
NRAOEPGRQVNRKLAKYKLGLEAGYSIDQIRTCORPTPIPVVAGNRDEQVEITR
VEAEVAKAFLAALGTTPANKPGRITMPLTTEHBDMDIRTKANALCR/LAHTVSLAY
AIEGMSRHYTFRPERASDCYSYRÖHGAQCSVLRSMAAAPLDDPSDFRA
VTÖGRILPQÖIETWRAPSDVLTINADRAAAGCMPLTÖWS"
7518. 8021
/gene="RSP0179"
/note="RS04691"
7518. 8021
/gene="RSP0179"
/function="miscellaneous; hypothetical/partial homology"
/note="Product confidence : hypothetical
Gene name confidence : hypothetical
predicted by Codon_usage
predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="HYPOTHETICAL PROTEIN"
/protein_id="CAD17330.1"
/db_xref="GI:17430647"
/translation="MELNAPDRMGOPTRAVRGAAIALSATALAPALADAKPGWQ
DEPARAAVEAÖQVEYVSGTHPLVLIHADPAMVAVQSDYFDVRRVAITPGSRILI
GKELRKVNARREIWIIGLQIPASGTLRLDPLQATMRGSGAGADLSGALGAMTS
EPLIVPH"

gene
8066. 8488
/gene="RSP0180"
/note="RS04692"
8066. 8488
/gene="RSP0180"
/function="miscellaneous; hypothetical/global homology"
/note="Product confidence : hypothetical
Gene name confidence : hypothetical
predicted by Codon_usage
predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="CONSERVED HYPOTHETICAL PROTEIN"
/protein_id="CAD17331.1"
/db_xref="GI:17430648"
/translation="MGANSHASFGNDADIDRAAEVATNDLSRVESAPDAVLAAGDD
LEAPASQAVAAAEVVARLQÖHGWVRNADSSIDAMIRRTGITPPROIYHQAQALDR
IVCEPSALLEMOSEDDCKRDVAIRELRALDGAAPAR"
complement(8552. 8956)
/gene="RSP0181"
/note="RS04693"

gene
8956. 8956
/gene="RSP0181"
/note="RS04693"

	fasta scores: E(): 0, 60.8% id in 227 aa, and to Burkholderia pseudomallei transcriptional activator Protein IIR SW:IKR_BURPS (O31395) (229 aa) fasta scores: E(): 0, 57.6% id in 224 aa"
	/codon_start=1 /transl_table=1 /product="transcriptional activator" /protein_id="CAC90814.1" /db_xref="GI:15980015"
Query Match	13.9%; Score 69.8; DB 1; Length 334050;
Best Local Similarity	51.8%; Pred. No. 0.00013;
Matches 158; Conservative	0; Mismatches 147; Indels 0; Gaps 0;
OY 62	accgcggacgagcactgcgccttcgccggcgcccccgcgaattgtcggagacgcgcacatc 121
Db 260521	AOCCTGATAGATATGATATTTTTGGCTCCTATTGTGGAAATGTCGACGCCCGATTGCC 260580
OY 122	cccaccgcgagactcaacggcgcccttgcttgcctgtctaactgagccgctgtgcgtg 181
Db 260581	CCATTAGCCCCCTGCACAGCGTGTTACTCATACAGCATATTATGCTATTGATATGG 260640
OY 182	gcgggtccgcgagatgcacatcgttgtcttcgtcatccagaagaatcgtgcgcgcgac 241
Db 260641	GCGATTATTGGAAAACTTGATATCCATGCTCCACCAACGGCAAAGATGTTGTGGCAA 260700
OY 242	cgtgtaaagctggttcagccgcctcagagtcggtgtgtgcgcgcacatcgtccgcgat 301
Db 260701	CATTCCAAAATTTATCCAACCCGCAAGACCCTGAGCTGGTGAAGAAATTATGTCGGAG 260760
OY 302	ggccaacacctcaaggcgccgcgcgagatctcgtacgtcagctgtcacatcgcgggtgtg 361
Db 260761	GGGGATACCGCTTGATGCGCGGACAGGTGCTATTGTCATTAAACCTATTTGCAATGATGCC 260820
OY 362	gagtgt 366
Db 260821	GAGGT 260825
RESULT 14	
BPCYADE	
LOCUS	BPCYADE 6645 bp DNA linear BCT 08-SEP-1992
DEFINITION	B.pertussis genes cyah, cyab, cyad and cyae.
ACCESSION	X14199
VERSION	X14199.1 GI:39731
KEYWORDS	adenylate cyclase, cya gene; cyA gene; cyab gene; cyad gene; cyAE gene; hemolysin; secreted protein; toxin.
SOURCE	Bordetella pertussis.
ORGANISM	Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae; Bordetella.
REFERENCE	1 (bases 1 to 2040)
AUTHORS	Glaser,P., Sakamoto,H., Bellalou,J., Ullmann,A. and Danchin,A.
TITLE	Secretion of cyclolysin, the calmodulin-sensitive adenylate cyclase-haemolysin bifunctional protein of Bordetella pertussis
JOURNAL	EMBO J. 7 (12), 3997-4004 (1988)
MEDLINE	89091151
REFERENCE	2 (bases 1 to 6645)
AUTHORS	Danchin,A.
TITLE	Direct Submission
JOURNAL	Submitted (08-SEP-1992) A. Danchin, Institut Pasteur, Unite de Regulation de l'Expression, Genetique, 28 rue du Docteur Roux, 75724 Paris Cedex 15, FRANCE
COMMENT	see Y00545 for upstream cya gene; cya operon is organized CYABDE, cyab (712 aa) is initiated 78 bp downstream of cyah stop; cyab stop overlaps with cyad initiation (440 aa).
FEATURES	location/Qualifiers
SOURCE	1..6645
	/organism="Bordetella pertussis"
	/strain="18323."
	/db_xref="taxon:520"
gene	1..725
	/gene="cyAH"

```

CDS
<1..658
/gene="cyaA"
/codon_start=2
/transl_table=11
/protein_id="CAA32411.1"
/db_xref="GI:39732"
/db_xref="SWISS-PROT:P15318"
/translation="PGRGLDAGAKGVFLSLGKGFASIMDEPETSNTLNENAVGSAR
DDVLIGAGAVNLGAGNDVLSGAGSDVLLSGDNDLSDGAGNDLFGCGGDDY
LFGVYGHDTYESGGHDTIRINAGADOLWFARQGNDEIRLIGTDDALVHMYRD
ADHVEITHANQAVDQAGIEKLEMAAQYDPGAAAAAPPAARVPPTLMGSLAVNMR"

RBS
721..725
/gene="cyaA"
736..2874
/gene="cyaB"
736..2874
/gene="cyaB"

CDS
/gene="cyaB"
/codon_start=1
/transl_table=11
/protein_id="CAA32412.1"
/db_xref="GI:39733"
/db_xref="SWISS-PROT:P18770"
/translation="MTSPVACASVPDGLCLVWLARYHGLAADPEQLRHEFAEQAF
CSRTQLAARVGLKVRHRRPAPAPLPAIALDRGGYVLYVPRREGADQAVL
IQRGQAPARLQGADEFELMAGELLCCASPTQALARPDSWFTIPALVHRHLIGE
VLLISLVQFISLTPLEFYVMDVNLNNAETLVITVGFLLAILEALITGRTY
VLAHTSKLDELGLARLYAHLRLPLAVQARVDSVARELHIAFLGNATV
LIDVYSVVFIAVMEFYSVKLTLVLAALPCYFLLSVLTPVLRRLHVKFERGAE
AFIVETVSGIDTVKSLAVEPQWRMDROLAGYVAGLSVANAMANTGYLLISRL
PREGSGMARGRARMTYGEIAFMLSGHVTPYIRLAOLMNDFOQTGVSMDRGDI
INCRTEVYAGKQAPALRGSTIELDRSPRYPPADALRNVLRIAPGSEITFNSVADI
SKSTLRLIQRMEYVADGRVILIDGHIDIGYDSASLRQGLVYLGESITFNSVADI
ALTRGASMEHVAARLGAHEFTCOLPEGDTMLGEGVLSGCGRORIGIAVALI
HRPVLILDEATSLADYSEHILIQRMNDICGRVILIIAHLASVRCADRIVMEGG
EVAECGSHETLLAAGLYARLQALQAGEAG"

RBS
2862..2866
/gene="cyaD"
2862..2866
/gene="cyaD"
2871..4193
/gene="cyaD"
2871..4193
/gene="cyaD"

CDS
/codon_start=1
/transl_table=11
/protein_id="CAA32413.1"
/db_xref="GI:39734"
/db_xref="SWISS-PROT:P11091"
/translation="MRRLRELAAHRGVVAAWROHRRRPAQWDPVTEFLPSALS
LOERPISPARMLARILMALAGALVWSVKTEIYVHAAGKVPVPGQSKIIAASFTG
RYARVIVADNSRYARAGVLDLADGVTAEERKWRFOAQAORDEARSAMIADLTG
RAPVLAEADPDGMAAOSYLDOSYADVOAOLRASTAIATARDVGLTOLAHARG
LRDGDVSOQATLEKQAKMTLEGRLOSEAORAAICOTTRQAEPTILARLAAOA
EQEIARTSAQRSLVLTAPVDGVQQLVLTSTVAAATOPLMVYVPSGAGIDVOAOL
DSKDIFVRAGAATVKGAVDTYTKYTGLEGKYLVSPTVVDROHSTRVITALLAH
PALEVKGKRLKEGMAYOVADIRFGSRRLIEYLLSPVARRAGESLGER"

gene
4184..4187
/gene="cyaE"
4184..4187
/gene="cyaE"
4195..5619
/gene="cyaE"
4195..5619
/gene="cyaE"

RBS
721..725
/gene="cyaA"
736..2874
/gene="cyaB"
736..2874
/gene="cyaB"

CDS
/gene="cyaB"
/codon_start=1
/transl_table=11
/protein_id="CAA32414.1"
/db_xref="GI:39735"
/db_xref="SWISS-PROT:P11092"
/translation="MAAVYRRGRALALMLMGFALSVGVARADGLATPPAFEGC
AAPAVSWPCPPADRLDDILALEALIDALCHSPALRGMARIKARSFEVGLARVYP
SVALSAGRSRQKRSTGLGDEGVNRNNAVYTLAMRLFDGSRASALRAQAQOLDEAQA

```

```

YGAVIDOKIAEVNGAYYEATAROLAHVEEDTEIARRSASIAARRACILSHGDVL
HAQALERARLQAQAGAAQARALAGIAOYVGNPATPIYVAPGLAPORDEBELAO
WLRDARORHPRATKAQGLAATNOVDYARATMTPTDLSIGHYNSSENVSYPAGSS
RSVSASLNLRIPLFDGPARHRIGGARAEOVORALIDQALATGAAYARAYADLRRA
RASHEASLRMLKAKARAAYESDLRRYEVGGVABLLRQSDWLSARORHVLTAQOLRT
RALALLAAGELGRSTIDGPPKE"

misc_feature
5706..>6645
/gene="cyaX"
/note="unknown ORF"

gene
5706..6645
/gene="cyaX"

gene
5760..6617
/gene="cyaX"
5760..6617
/gene="cyaX"

CDS
/note="unknown ORF"
/codon_start=1
/transl_table=11
/protein_id="CAA32415.1"
/db_xref="GI:39736"
/db_xref="SPTREMBL:Q45360"
/translation="MRPVHDCGRAAPCVAVGROPAPVAGDLGRHASDPGLPPTPD
GAGEQFRSAAAEIYKRYIDARGEGAGDHDVNLVAMPDCLATSRFPAMCARMOVG
GAASVRLHGVNVDSEVNLAVGVNDLIVCFRHAVEDIOLDQOSYMRACVADVRLYR
PAGMALDEAEIACEAPQARYPYVTSAGATYRRLIELSCGPLORADQRCVETFPD
GAPHLVRAGVGLAMLPDSSARAAVEHGIYVADRCAIPMETCYARRHGGQEAIVE
RVFDQFGNG"

BASE COUNT      956 a      2142 c      2441 g      1106 t
ORIGIN

Query Match      12.5%; Score 62.6; DB 1; Length 6645;
Best local Similarity 50.5%; Pred. No. 0.006;
Matches 152; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

QY      65  cggagagagctgcttctctgcccggccctcgaaatgtcgagacgcgcgcacatccc 124
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      2967  CCGGTGGAGACCGAGTTCCTCCCTTCGGCTTCCGTGACGAGCGTCGATATGCC 3026

QY      125  acccgagacccacgcgcgcctgtctgtcttctactgcgcgctgycgtggcg 184
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      3027  ACGGCCGTTGGCTGGCGCGGCACATACGATGCGCGTGGCGCGCGCTGCTG 3086

QY      185  ggtctggcagagatcgacatcgctgtcttcgcacacgaagaatcgycggcgagccgt 244
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      3087  GTAGTGGCGCAAGACCGGATCGTCGTGCATGCGGCGGCAAGGTGCGCGGCGGACG 3146

QY      245  gtaagctggttagccgcgcgcaggtcgctggtgctggtgcacatcgtcccgatgyc 304
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      3147  AGCAAGATCATCGCGGCGTCGGAACCGGTCGCGCGCGCGCGCGCGGACGACAT 3206

QY      305  caaacgtcaagcgcgcgagatctgatacagctgatacctatcgcggtgtgtgat 364
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      3207  AGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGAA 3266

QY      365  g 365
DB      3267  G 3267

RESULT 15
AF193064      7493 bp      DNA      linear      BCT 14-NOV-1999
LOCUS      AF193064
DEFINITION      Caulobacter crescentus truncated S-layer protein (rsaA), ABC
translocator (rsad), and membrane forming unit (rsaf) genes,
complete cds.
ACCESSION      AF193064
VERSION      AF193064.1
KEYWORDS      GI:6409374
SOURCE      Caulobacter crescentus.
ORGANISM      Caulobacter crescentus.
Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
Caulobacter.

```


Tue Jul 16 05:46:21 2002

us-09-484-577a-3.rge

Page 18

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 15, 2002, 22:20:07 ; Search time 1623.58 Seconds
(without alignments)
4164.855 Million cell updates/sec

Title: US-09-484-577A-3
Perfect score: 501
Sequence: 1 actccagcctctaccga.....cagaatgcgcgatgatcat 501

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues
1 number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estha:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_ges:*
13: em_gss_hum:*
14: em_gss_hiv:*
15: em_gss_pin:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60.6	12.1	932	12	CNS0070E
2	55	11.0	925	12	CNS0091P
3	44.8	8.9	508	9	AV630811
4	43.6	8.7	772	12	AF075872
5	42	8.4	839	12	CNS004NB
6	40.8	8.1	932	12	CNS0072Q
7	40.4	8.1	774	10	BI956591
8	40.2	8.0	581	9	AV921113
9	40.2	8.0	586	9	AV913605
10	40.2	8.0	605	9	AV918908
11	40.2	8.0	659	9	AV931532
12	40.2	8.0	660	9	AV946608
13	39.8	7.9	639	9	AI770858
14	39.6	7.9	574	10	BM349363
15	39.6	7.9	576	10	AI881920
16	39.6	7.9	577	10	BM267007
17	39.6	7.9	645	12	CNS01213

Result No.	Score	Query Match	Length	ID	Description
18	39.6	7.9	686	10	BI959511
19	39.6	7.9	692	12	CNS007MH
20	39.6	7.9	754	10	BF859678
21	39.4	7.9	814	10	BG310235
22	39.2	7.8	940	12	CNS0625X
23	39.2	7.8	870	10	BI950671
24	38.8	7.7	446	10	BM372629
25	38.8	7.7	603	10	BF253745
26	38.8	7.7	660	9	BE215392
27	38.6	7.7	844	12	CNS0052P
28	38.4	7.7	642	10	BI954172
29	38.4	7.7	925	12	CNS0091P
30	38	7.6	265	10	R01662
31	38	7.6	755	10	BF616985
32	37.8	7.5	672	9	BB528893
33	37.5	7.5	424	10	BE404013
34	37.4	7.5	455	9	BB833701
35	37.4	7.5	698	9	BB610391
36	37	7.4	392	9	AV644920
37	37	7.4	393	9	AV644931
38	37	7.4	424	10	BE497481
39	37	7.4	539	9	AV640485
40	37	7.4	629	9	AV912950
41	37	7.4	667	12	A2131633
42	36.8	7.3	441	10	BE604637
43	36.8	7.3	510	10	BE585055
44	36.8	7.3	521	9	AV434008
45	36.8	7.3	600	9	AV602736

ALIGNMENTS

RESULT 1
CNS0070E/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence JET3 end of BAC # BAC1421 of RPCI-98 library from Drosophila melanogaster (fruit fly) genomic survey sequence.

ACCESSION
AL066254.1 GI:4945121
VERSION
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster
fruit fly.
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage; BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr)
- Web: www.genoscope.cns.fr
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammosser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source
1..932
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCI-98"

[illegible]

CRS0072Q

LOCUS	932 bp	DNA	linear	GSS 03-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence T7 end of BAC #			
	BACR14B09 of RPC1-98 library from Drosophila melanogaster (fruit			
	fly), genomic survey sequence.			
	1507732			

ACCESSION	AL006742	GI:4945205
VERSION	AL066742.1	
KEYWORDS	GSS.	
SOURCE	fruit fly.	
ORGANISM	Drosophila melanogaster	

REFERENCE
1 (bases 1 to 932)

AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage

COMMENT
BP 191 91006 EVRY cedex - FRANCE (E-mail : segrete@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
Determination of this BAC-end sequence was carried out as part of a

collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the *Drosophila melanogaster* genome using these BACs. For further information please see http://www.fruitfly.org/TheBDGP/Drosophila_melanogaster_BAC_library was prepared by Kazuhiro Oosegawa and Aaron Mammoso in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
Source

BASE COUNT	155 a	202 c	241 g	91 t	243 others
ORIGIN	/organism="Drosophila melanogaster" /db_xref="taxon:7227" /clone_lib="RPCT-98" /clone="BACR14B09" /note="end : 77"				

Query Match	8.1%	Score 40.8;	DB 12;	Length 932;
Best Local Similarity	30.6%	Pred. No. 13;		
Matches	96;	Conservative	64;	Mismatches 154;
			Indels	0;
			Gaps	0;

Qy *151 tgcgacctgtctactacgcccgttgcggtgcgagtcctgcgcgaagtcatcattcttgc 210
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 586 BSBFKYTBKAGGCGGCGSSGCCSBSGSCRGGGGCGCGSGSGCCGCCCGGSSBSCGCSCGC 645

Oy 211 ttctgcataccgaaagatctgtgcacggagacactgtgaagcttgattcaacgcgtcagagt 270
 Db 646 GSGGCGCCGCCCCSSGSGSGSGSGSGSGSGCCCGGCGCCSSCGCCSSCGSCCCCGCG 705
 Oy 271 cgagcgtgtgtcggagccatcatgtccgagatgagccaaacggtcaagcgcgagattct 330
 Db 706 CGSGCGSGCGCGCGCSGSGSCGSCGCCSCGCCSCGCCSCGSGSGSGSGSGSGSG 765
 Oy 331 gatcgaagcttgatcatcttcgcgaggtgtgtgtgtgattgtccgcccgtcagaagttcatac 390
 Db 766 CGSGCCCGSGSGSGSGCGCGCGCGSGSGSGSGSGSGSGSGCGCCSSSGSGSGSGSGCS 825
 Oy 391 ggtgtcgtgcgcgcccaagatgcacacatcttgcagactttcttcaacgaagattca 450
 Db 826 GGSAGCGCSGSGGCGCCSSGCGGCGCGCSGCGSGGCGCGSGCGCCCGCGCGCGCG 885
 Oy 451 ccgcgcagttgcg 464
 Db 886 GSGSGCGGSGCGSG 899

RESULT 7
BI956591
LOCUS

LOCUS	BIOSOURCE	174 bp	MINI	LINEAL	ESI 22 OCT 2001
DEFINITION	HVSMEN0004F02f Hordeum vulgare rachis EST library HVCDA0015				

ACCESSION (normal) Hordeum vulgare cDNA clone HY5MENU04F0U1, mRNA sequence.
VERSION B1956591
KEYWORDS B1956591.1 GI:16307844
SOURCE EST.
barley.

ORGANISM *Hordeum vulgare*
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
; Triticeae; *Hordeum*.

REFERENCE
1 (bases 1 to 774)
Wing, R., close, T.J., Kleinhofs, A., Wise, R., Chin, A., Begum, D.,
Frisc, D., Atkins, M., Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons

TITLE
'J', Oates, R. and Main, D.
Development of a genetically and physically anchored EST resource
for barley genomics: Morex rachis cDNA library

JOURNAL
 COMMENT
 Unpublished (2001)
 Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Total hg bases = 534
 Seq primer: AATTACCCCTCAGCTAAAGG
 High quality sequence stop: 686.
 location/qualifiers

SOURCE

```

/organism="Hordeum vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"
/clone="HWSM0004F02f"
/clone_id="Hordeum vulgare rachis EST library HVCN0015
(normal)"
/tissue_type="Rachis"
/lab_host="TJc121"
/note="Vector: pBluescript SK(-); Site1: EcoRI; Site2:
XhoI; Plants were grown at Washington State University,
Pullman, WA in a greenhouse, the rachises were excised and
frozen in liquid nitrogen (Kleinof's lab). In the TJ Close
lab at the University of California, Riverside total RNA
was prepared, poly(A) was purified, one primary
unamplified cDNA library was made, and 1 million pfu were
in vivo packaged to give pBluescript SK(-) cDNA phagemids
(Chin). Phagemids were plated and picked at the Clemson
University Genomics Institute (CUGI) (Begum, Palmer,
Frisch, Atkins and Windy). Plasmid DNA preparations, DNA

```

```

/organism="Hordeum vulgare subsp. vulgare"
/cultivar="Haruna Nijo"
/db_xref="taxon:112509"
/clone="bags1a15"
/clone_lib="K. Sato unpublished cDNA library, cv. Haruna

```

BASE COUNT	ORIGIN
134 a	162 c 153 g 137 t

Query Match 8.0%; Score 40.2; DB 9; Length 586;
 Best Local Similarity 45.2%; Pred. No. 15;
 Matches 147; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

QY 88 ggcgcctcgaattgtcgcagagcgccgcatctccacgcgagactcaagcgccctt 147
 DB 553 GGACGCCGACGTCATCACCATTGAGAACTCAGCTCCAGAGAGCTCTCTCGGTCTT 494
 Db 4 GGACGCCGACGTCATCACCATTGAGAACTCAGCTCCAGAGAGCTCTCTCGGTCTT 63
 QY 148 gcttcgtcctgttctactcgtcgcggtgtggtggtggtcgcgagagatcagatcgt 207
 DB 64 CCGCAGAGGCGGTGACCTTACGAGGCGGCGCATTTGGCCCGCTTACGACATCCTCC 123
 QY 208 tgccttcgcacagaaagatcgtgcgcggtggtggtggttgaagctgtgtcagcgctga 267
 DB 124 CAGGATCCCTCCACGAGAGGAGATTGCGACCGCTCAACAAGATGCTCGCGTCTCA 183
 QY 268 gtcggtgtgtgtggtggtggtcgcacatctgtcgcgtggtggtggtggtggtggt 327
 DB 184 CACCAACATCTCTCTGCTGAGACCTGACTGCGGCTCAAGACCGCAAGTACACCGAGT 243
 QY 328 tctgacagctggtatccatctcgcggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 387
 DB 244 GAAGCCCGCCCTGACTAACATGTCGTCGCGCAAGCTCATCCGACCCAGCTCGCAG 303
 QY 388 cagcgtgtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 412
 DB 304 CACGAAGTGAGGAGCCATGAATCG 328

RESULT 10

AV918908/c 605 bp mRNA linear EST 18-JAN-2002
 LOCUS AV918908 K. Sato unpublished cDNA library, cv. Haruna NiJo
 DEFINITION germination shoots Hordeum vulgare subsp. vulgare cDNA clone
 AV918908
 AV918908
 accession AV918908.1 GI:18214687

KEYWORDS

EST.
 Hordeum vulgare subsp. vulgare.
 SOURCE Hordeum vulgare subsp. vulgare.
 ORGANISM Hordeum vulgare subsp. vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
 ; Triticeae; Hordeum.
 1 (bases 1 to 605)

REFERENCE Sato, K., Saitoh, D. and Takeda, K.
 AUTHORS Barley EST sequencing project in NIG and Okayama Univ
 TITLE Unpublished (2002)
 JOURNAL Contact: Tadasu Shin-1
 COMMENT Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp.

FEATURES

source 1. 605
 /organism="Hordeum vulgare subsp. vulgare"
 /cultivar="Haruna NiJo"
 /db_xref="taxon:112509"
 /clone="basd22k19"
 /clone_lib="K. Sato unpublished cDNA library, cv. Haruna
 NiJo germination shoots"
 /tissue_type="shoots"
 /dev_stage="germination"
 BASE COUNT 140 a 152 c 179 g 133 t 1 others
 ORIGIN

Query Match 8.0%; Score 40.2; DB 9; Length 605;
 Best Local Similarity 45.2%; Pred. No. 15;
 Matches 147; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

QY 88 ggcgcctcgaattgtcgcagagcgccgcatctccacgcgagactcaagcgccctt 147
 DB 553 GGACGCCGACGTCATCACCATTGAGAACTCAGCTCCAGAGAGCTCTCTCGGTCTT 494
 Db 4 GGACGCCGACGTCATCACCATTGAGAACTCAGCTCCAGAGAGCTCTCTCGGTCTT 63
 QY 148 gcttcgtcctgttctactcgtcgcggtgtggtggtggtcgcgagagatcagatcgt 207
 DB 493 CCGCAGAGGCGGTGACCTTACGAGGCGGCGCATTTGGCCCGCTTACGACATCCTCC 434
 QY 208 tgccttcgcacagaaagatcgtgcgcggtggtggtggttgaagctgtgtcagcgctga 267
 DB 433 CAGGATCCCTCCACGAGAGGAGATTGCGACCGCTCAACAAGATGCTCGCGTCTCA 374
 QY 268 gtcggtgtgtgtggtggtggtcgcacatctgtcgcgtggtggtggtggtggtggt 327
 DB 373 CACCAACATCTCTCTGCTGAGACCTGACTGCGGCTCAAGACCGCAAGTACACCGAGT 314
 QY 328 tctgacagctggtatccatctcgcggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 387
 DB 313 GAAGCCCGCCCTGACTAACATGTCGTCGCGCAAGCTCATCCGACCCAGCTCGCAG 254
 QY 388 cagcgtgtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 412
 DB 253 CACGAAGTGAGGAGCCATGAATCG 229

RESULT 11

AV91532/c 659 bp mRNA linear EST 18-JAN-2002
 LOCUS AV91532 K. Sato unpublished cDNA library, cv. Haruna NiJo second
 DEFINITION leaf stage seedling leaves Hordeum vulgare subsp. vulgare cDNA
 clone basd23c10 3', mRNA sequence.
 AV91532
 AV91532
 accession AV91532.1 GI:18227329

KEYWORDS

EST.
 Hordeum vulgare subsp. vulgare.
 SOURCE Hordeum vulgare subsp. vulgare.
 ORGANISM Hordeum vulgare subsp. vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
 ; Triticeae; Hordeum.
 1 (bases 1 to 659)

REFERENCE Sato, K., Saitoh, D. and Takeda, K.
 AUTHORS Barley EST sequencing project in NIG and Okayama Univ
 TITLE Unpublished (2002)
 JOURNAL Contact: Tadasu Shin-1
 COMMENT Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp.

FEATURES

source 1. 659
 /organism="Hordeum vulgare subsp. vulgare"
 /cultivar="Haruna NiJo"
 /db_xref="taxon:112509"
 /clone="basd23c10"
 /clone_lib="K. Sato unpublished cDNA library, cv. Haruna
 NiJo second leaf stage seedling leaves"
 /tissue_type="seedling leaves"
 /dev_stage="second leaf stage"
 BASE COUNT 152 a 159 c 197 g 149 t 2 others
 ORIGIN

Query Match 8.0%; Score 40.2; DB 9; Length 659;
 Best Local Similarity 45.2%; Pred. No. 16;
 Matches 147; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

QY 88 ggcgcctcgaattgtcgcagagcgccgcatctccacgcgagactcaagcgccctt 147
 DB 552 GGACGCCGACGTCATCACCATTGAGAACTCAGCTCCAGAGAGCTCTCTCGGTCTT 493

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 15, 2002, 23:19:32 ; Search time 205.98 seconds

(without alignments)
4176.008 Million cell updates/sec

Title: US-09-484-577A-3

Perfect score: 501

Sequence: 1 actctcagcctctcaccga.....cagaatgcgcgatcatcat 501

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 segs, 858457221 residues 3472872

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N.Geneseq_032802:*

1: /SIDSI/gcgdata/hold-geneseq/geneseg-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/hold-geneseq/geneseg-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/hold-geneseq/geneseg-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/hold-geneseq/geneseg-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/hold-geneseq/geneseg-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/hold-geneseq/geneseg-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/hold-geneseq/geneseg-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/hold-geneseq/geneseg-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/hold-geneseq/geneseg-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/hold-geneseq/geneseg-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/hold-geneseq/geneseg-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/hold-geneseq/geneseg-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/hold-geneseq/geneseg-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/hold-geneseq/geneseg-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/hold-geneseq/geneseg-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/hold-geneseq/geneseg-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/hold-geneseq/geneseg-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/hold-geneseq/geneseg-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/hold-geneseq/geneseg-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/hold-geneseq/geneseg-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/hold-geneseq/geneseg-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/hold-geneseq/geneseg-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/hold-geneseq/geneseg-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/hold-geneseq/geneseg-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
C 1	88.8	17.7	28804 17	AAT37329 Sphingian biosynthe
C 2	88.8	17.7	28804 18	AAT92474 Sphingomonas genus
C 3	88.8	17.7	28804 20	AAT98812 Sphingomonas S88 s
C 4	88.8	17.7	28804 20	AAV81474 Chrtomosomal fragme
C 5	86	17.2	1305 21	AAZ54336 Neisseria meningit
C 6	86	17.2	1428 21	AAZ54335 Neisseria meningit
C 7	86	17.2	24158 21	AAH15132 N. meningitidis pa
C 8	86	17.2	349980 21	AAF21611 Neisseria meningit
C 9	86	17.2	349980 21	AAF21612 Neisseria meningit

C 10	86	17.2	837096 21	AAH81489 N. meningitidis pa
C 11	51.6	10.3	5120 18	AAT73218 AprIBD gene. Acti
C 12	51.6	10.3	5120 21	AAZ88585 A. pleuropneumonia
C 13	47.8	9.5	6400 20	AAZ22701 Pseudomonas fluore
C 14	47	9.4	4403765 22	AAI99683 Mycobacterium tube
C 15	47	9.4	4411529 22	AAI99682 Mycobacterium tube
C 16	45	9.0	7184 15	AAO70050 Mycobacterium tube
C 17	43.4	8.7	7183 11	AAO60774 pit gene of Pasteu
C 18	43	8.6	24379 18	AAT93095 Sequence encoding
C 19	43	8.6	24379 19	AAV25925 Streptomyces iteno
C 20	39	7.8	1761 22	AAH66404 Streptomyces roseo
C 21	39	7.8	1998 22	AAAF72013 C glutamicum codin
C 22	39	7.8	349980 22	AAH68528 Corynebacterium gl
C 23	37.6	7.5	1553 23	AAH89358 C glutamicum codin
C 24	37.6	7.5	2931 23	AAH92854 DNA encoding novel
C 25	37.6	7.5	10732 21	AAAI0594 Gene encoding novel
C 26	37.6	7.5	4403765 22	AAI99683 Mycobacterium tube
C 27	37.2	7.4	1194 17	AAT58555 Streptomyces prist
C 28	37.2	7.4	4496 17	AAT58553 Streptomyces prist
C 29	37.2	7.4	58857 21	AAH58471 Nucleotide sequenc
C 30	37	7.4	1951 19	AAV11459 C. acidivorans gam
C 31	37	7.4	1981 21	AAZ50482 Corn sulplate perm
C 32	36.4	7.3	1739 22	AAH19384 Rat CCAAT/enhancer
C 33	36.2	7.2	16020 21	AAA39283 Streptomyces nogal
C 34	36	7.2	1136 21	AAAC47742 zea mays DNA fragm
C 35	35.8	7.1	836 19	AAH62157 HSY-2 strain SB5 C
C 36	35.8	7.1	3765 21	AAH55782 Nucleotide sequenc
C 37	35.8	7.1	53500 21	AAH55842 Complete nucleotid
C 38	35.8	7.1	117213 19	AAH62176 HSY-2 strain SB5 C
C 39	35.8	7.1	154746 24	AAH25519 Human herpesvirus
C 40	35.6	7.1	4411529 22	AAI99682 Mycobacterium tube
C 41	35.4	7.1	48300 22	AAH61281 Mycobacterium tube
C 42	35	7.0	1119 24	AAH24963 N. magdarii bacter
C 43	35	7.0	2025 21	AAH49036 Human G-protein co
C 44	34.8	6.9	1470 21	AAH38391 DNA encoding a hig
C 45	34.8	6.9	1500 22	AAH19383 Pseudomonas sp. WF

ALIGNMENTS

RESULT 1	AAAT37329/C	strand; DNA; 28804 BP.
ID	AAAT37329	standard; DNA; 28804 BP.
AC	AAAT37329;	
DT	30-NOV-1996	(first entry)
DE	Sphingian biosynthetic gene region.	
XX	Sphingian; polysaccharide; spsb gene; glucosyl-IP-transferase; ds.	
XX	Sphingomonas strain S88 (ATCC 31554).	
OS		
XX		
FH	Key	Location/Qualifiers
FT	CDS	complement (1942..1944)
FT		/*tag= a
FT		/codon_start= 1942..1944
FT	CDS	/note= "spsg gene putative initiation codon"
FT		complement (3311..3313)
FT		/*tag= b
FT		/codon_start= 3311..3313
FT		/note= "spsg gene putative initiation codon"
FT	CDS	complement (5323..5325)
FT		/*tag= c
FT		/codon_start= 5323..5325
FT		/note= "spsr gene putative initiation codon"
FT	CDS	5526..5528
FT		/*tag= d
FT		/codon_start= 5526..5528
FT		/note= "sps0 gene putative initiation codon"
FT	CDS	complement (7076..7078)

PT		/tag= e	/codon_start= 7076..7078	
FT		/note= "spst gene putative initiation codon"		
FT	CDS	7588..7590		
FT		/tag= f		
FT		/codon_start= 7588..7590		
FT	CDS	/note= "spsk gene putative initiation codon"		
FT		8643..8645		
FT		/tag= g		
FT		/codon_start= 8643..8645		
FT	CDS	/note= "pspl gene putative initiation codon"		
FT		complement (10938..10940)		
FT		/tag= h		
FT		/codon_start= 10938..10940		
FT	CDS	/note= "spsj gene putative initiation codon"		
FT		11569..11571		
FT		/tag= i		
FT		/codon_start= 11569..11571		
FT		/note= "spsf gene putative initiation codon"		
FT	CDS	12886..12888		
FT		/tag= j		
FT		/codon_start= 12886..12888		
FT		/note= "spsd gene putative initiation codon"		
FT		15165..15167		
FT		/tag= k		
FT	CDS	/codon_start= 15165..15167		
FT		/note= "spse gene putative initiation codon"		
FT	CDS	15883..15885		
FT		/tag= l		
FT		/codon_start= 15883..15885		
FT		/note= "urf12 gene putative initiation codon"		
FT	CDS	16749..16750		
FT		/tag= m		
FT		/codon_start= 16748..16750		
FT		/note= "urf26 gene putative initiation codon"		
FT	CDS	complement (21082..21084)		
FT		/tag= n		
FT		/codon_start= 21082..21084		
FT	CDS	/note= "atrb gene putative initiation codon"		
FT		21706..23118		
FT		/tag= o		
FT		/label= spsb		
FT		/product= glucosyl IP-transferase		
FT	CDS	23238..23240		
FT		/tag= p		
FT		/codon_start= 23238..23240		
FT		/note= "rhsa gene putative initiation codon"		
FT	CDS	24113..24115		
FT		/tag= q		
FT		/codon_start= 24113..24115		
FT		/note= "rhsc gene putative initiation codon"		
FT		24683..24685		
FT		/tag= r		
FT		/codon_start= 24683..24685		
FT	CDS	/note= "rhsh gene putative initiation codon"		
FT		25744..25746		
FT		/tag= s		
FT		/codon_start= 25744..25746		
FT		/note= "rhsl gene putative initiation codon"		
FT	CDS	complement (27534..27536)		
FT		/tag= t		
FT		/codon_start= 27534..27536		
FT		/note= "urfi1 gene putative initiation codon"		
FT	CDS	27747		
FT		/tag= u		
FT		/codon_start= 27747..27749		
FT		/note= "urif34 gene putative initiation codon"		
PN	EP728841-A2.			
XX	28-AUG-1996.			
PD				
XX	24-JAN-1996;	96EB-0300467		

XX 24-JAN-1995; 95US-0377440.
PR
PA (SHIN-) SHINETSU BIO. INC.
PA (SHIE) SHINETSU CHEM CO LTD.
XX
PI Armentrout RW, Mikolajczak M, Pollock TJ, Thorne L;
PI Yamazaki M;
XX
XX WPI: 1996-386292/39.
DR P-PSDB: AAW03997.
XX
XX New isolated DNA from *Sphingomonas* sp. - used for transforming
PT recipient bacteria to obtain hyper-producers of sphingan
PT polysaccharide(s).
XX
XX Claim 32; Page 56-70; 105pp; English.
PS
XX A 28.8 kb chromosomal fragment of *Sphingomonas* strain S88 was
CC isolated on the basis of its ability to restore sphingan
CC biosynthetic capability to *Sphingomonas* mutant S88m260. It
CC contains 23-25 genes, including sps genes coding for biosynthesis of
CC the polysaccharide sphingan, this genes coding for dtdp-(L)thamose
CC biosynthesis, atpB genes coding for a transport function and some
CC unidentified open translation reading frames (url). The spsB gene
CC was identified that is believed to code for glucosyl IP-transferase
CC (AAW03997), an enzyme catalyzing the first step of assembly of
CC sphingan carbohydrates. DNA fragments of S88 can be inserted into
CC a vector in multiple copies and used to produce engineered bacteria
CC that are hyper-producers of sphingan.
XX
XX Sequence 28804 BP; 4974 A; 9806 C; 9228 G; 4796 T; 0 other;

Query Match	Similarity	17.7%	Score 88.8	DB 17	Length 28804
Best Local	Similarity	56.5%	Pred. No. 6.2e-13		
Matches 165	Conservative	0	Mismatches 127	Indels	Gaps
QY	80	ttccgcgcgcgcgcgcgcgcgaattg	tcgagacgcccgcattcccaacgcgagactcacg	139	
DB	18808	TTCCGTCCCGCGCGCGCTCTGAGAT	TCATCGAACGTCGGGTGTGCGCCACCGCGGCTCAACC	18749	
QY	140	gccgcctgtgctgtgctgtctgttct	ctactcgtgcgcgtgtggtgtgcgtgtctgcgcaggtac	199	
DB	18748	GGCCCGCGGTGATGTTGGCGGGGCT	GTGGCGATCACACACCGCTGTGGCAATGGCGCGCTG	18689	
QY	200	gacatcattgtctctgtgcataccaga	aaagatcgtgcgcgcgcgcgcgtgtaagctgtgttcag	259	
DB	18688	GAAATGTTGTGGCCCCACGACGAGG	CGCCGTCATGCCCGCGATGGCGAAGACCAAGATCTGTGCAG	18629	
QY	260	ccgcctcgcggtgcgctgtgtgtgcgc	acatactgtctcgagatgtgcacaaacgltcaagggc	319	
DB	18628	TCGCCCGAAGGCGGATGCTCCGGGCG	CAATTCTGGTGGCGGAGGGGCGAAGAGTTTCAGAGG	18569	
QY	320	ggcgcgagatctgatacgagcttgatc	atccatcgcgcggtgtgtgtgtgtgtgcg	371	
DB	18568	GGCCGAGGTGCTGATCAGCTCGATCC	ACCAATGTGGGAGCCGAGGCCCGCGC	18517	
RESULT 2					
AAT92474/C					
ID	AAT92474	standard	DNA	28804	BP.
XX					
AC	AAT92474				
XX					
DT	04-FEB-1998	(first entry)			
XX					
DE	Sphingomonas	genus	microbe	isolated	DNA sequence producing sphingian.
XX					
KM	Sphingomonas	microbe	sphingian	polysaccharide	biosynthesis gene;
XX					
KX	sphingian	S-88	spss	gene	ss.
OS	Sphingomonas	sp.			

CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent

PI Rappuoli R, Pizza M;



1
2
3
4

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 15, 2002, 22:46:27 ; Search time 43.46 Seconds
(without alignments)
2831.624 Million cell updates/sec

Title: US-09-484-577A-3
Perfect score: 501
Sequence: 1 actctccagctcaccga.....cagaatgcygcatgatcat 501

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues
Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*

1: /cgn2_6/ptodata/2/1na/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/1na/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/1na/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/1na/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/1na/PTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/1na/Dackfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	88.8	17.7	28804	2	US-08-592-874-1
C 2	88.8	17.7	28804	3	US-09-096-942-2
C 3	88.8	17.7	28804	3	US-09-096-867-2
C 4	51.6	10.3	5120	3	US-08-772-270A-6
C 5	51.6	10.3	8370	2	US-08-488-706-1
C 6	47	9.4	4403765	4	US-09-103-840A-2
C 7	47	9.4	4411529	4	US-09-103-840A-1
C 8	37.6	7.5	4403765	4	US-09-103-840A-2
C 9	37.2	7.4	1194	4	US-08-765-907A-9
C 10	37.2	7.4	4496	4	US-08-922-865-1
C 11	37	7.4	1951	4	US-08-922-865-1
C 12	35.6	7.1	4411529	4	US-09-103-840A-1
C 13	34.8	6.9	1500	4	US-09-593-711A-10
C 14	34.6	6.9	8438	1	US-07-945-283-1
C 15	34.4	6.9	1879	2	US-08-403-852D-5
C 16	34.4	6.9	1879	4	US-08-510-646B-5
C 17	34.4	6.9	1879	4	US-09-231-818-5
C 18	34.2	6.8	10095	3	US-08-822-586-45
C 19	34	6.8	1758	1	US-08-078-222B-1
C 20	34	6.8	1758	2	US-08-661-330A-1
C 21	34	6.8	1758	3	US-09-038-217A-1
C 22	34	6.8	49272	1	US-08-614-770A-1
C 23	33.8	6.7	933	4	US-09-105-390-43
C 24	33.8	6.7	1008	4	US-09-105-390-59
C 25	33.8	6.7	2810	4	US-09-105-390-6
C 26	33.2	6.6	2293	4	US-08-604-913B-12
C 27	33.2	6.6	2787	4	US-09-105-537-40

C 28	33.2	6.6	3004	1	US-08-276-213-6	Sequence 6, Appl
C 29	33.2	6.6	5970	3	US-09-320-878-21	Sequence 21, Appl
C 30	33	6.6	789	4	US-09-164-193-6	Sequence 6, Appl
C 31	33	6.6	1016	4	US-09-164-193-4	Sequence 4, Appl
C 32	32.8	6.5	5247	1	US-08-920-812-15	Sequence 15, Appl
C 33	32.8	6.5	5247	1	US-08-920-827-15	Sequence 15, Appl
C 34	32.8	6.5	5247	1	US-08-921-177-15	Sequence 15, Appl
C 35	32.8	6.5	5247	1	US-08-362-577C-15	Sequence 15, Appl
C 36	32.8	6.5	5247	2	US-08-920-828-15	Sequence 15, Appl
C 37	32.6	6.5	3576	4	US-08-976-259-79	Sequence 79, Appl
C 38	32.6	6.5	23673	4	US-09-773-816-1	Sequence 129, App
C 39	32.4	6.5	3065	4	US-08-599-895-1	Sequence 1, Appl
C 40	32.4	6.5	3877	2	US-08-211-290-1	Sequence 1, Appl
C 41	32.4	6.5	3877	3	US-09-322-676-1	Sequence 1, Appl
C 42	32.4	6.5	3877	4	US-09-466-036A-1	Sequence 1, Appl
C 43	32.4	6.5	3877	4	US-09-082-092-2	Sequence 2, Appl
C 44	32.2	6.4	295	4	US-08-998-416-662	Sequence 662, App
C 45	32.2	6.4	720	4	US-08-998-416-662	Sequence 662, App

ALIGNMENTS

RESULT 1
US-08-592-874-1/C
Sequence 1, Application US/08592874
Patent No. 5854034
GENERAL INFORMATION:
APPLICANT: POLLOCK, THOMAS J.
APPLICANT: YAMAZAKI, MOTOHIDE
APPLICANT: THORNE, LINDA
APPLICANT: MIKOLAJCZAK, MARCIA
APPLICANT: ARENTROUT, RICHARD W.
TITLE OF INVENTION: DNA SEGMENTS AND METHODS FOR INCREASING
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: JULES E. GOLDBERG
STREET: 261 MADISON AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016-2391
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,874
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/377,440
FILING DATE: 24-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: GOLDBERG, JULES E.
REGISTRATION NUMBER: 24,408
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-986-4090
TELEFAX: 212-818-9479
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 28804 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FRAGMENT TYPE: N-terminal
US-08-592-874-1
Query Match 17.7%; Score 88.8; DB 2; Length 28804;

[illegible]

```

RESULT 10
US-08-765-907A-6
; Sequence 6, Application US/08765907A
; Patent No. 6352839
; GENERAL INFORMATION:
; APPLICANT: BLANC, Veronique
; APPLICANT: THIBAUT, Denis
; APPLICANT: BAMA-JACQUES, Nathalie
; APPLICANT: BLANCHE, Francis
; APPLICANT: COUZET, Joel
; APPLICANT: BARRIERE, Jean-Claude
; APPLICANT: DEBUSSCHE, Laurent
; APPLICANT: FAMECHON, Alain
; APPLICANT: PARIS, Jean-Marc
; APPLICANT: DUTROC-ROSETT, Gilles
; TITLE OF INVENTION: Streptogramins And Method For Preparing Same By
; TITLE OF INVENTION: Mutasynthesis
; FILE REFERENCE: Streptogramin genes
; CURRENT APPLICATION NUMBER: US/08/765,907A
; CURRENT FILING DATE: 1997-03-20
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 4496
; TYPE: DNA
; ORGANISM: Streptomyces pristinaespiralis
;S-08-765-907A-6

```

[illegible][illegible]

```

RESULT 11
US-08-922-865-1/c
: Sequence 1, Application US/08922865
: Patent No. 6090616
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: MICROORGANISM, LACTAMASE ENZYME OBTAINED
: TITLE OF INVENTION: THEREFROM, AND THEIR USE
: NUMBER OF SEQUENCES: 2
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/922,865
: FILING DATE:
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1951 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Commamonas acidovorans
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 49..1773
: US-08-922-865-1

```

```

Query Match Similarity      7.4%: Score 37; DB 3; Length 1951;
Best Local Similarity      53.9%: Pred. No. 0.43;
Matches 76; Conservative 0; Mismatches 65; Indels 0; Gaps 0.

QY 228 tcgtgcggagcgacccgttlaaagctcgtgttcacgcgcctcgaggttcggcgtgtgcgggcca 287
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 434 TCGTGGCCCGGGCCGCGCATGTCCAGCAGGTTCACACAGCAGAGTCCGGCGGGGTGCGGGCCC 375

QY 288 ctcatgttcgcgagtggccaacgcgtcaagagccgcgagattctgatcgagctgtgacat 347
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 374 TTCAAGGCCACAGGGGGCCGCGACGAGAAAGTGAGCGGTGAGACAGATCCAGTCGCGCAGTCT 315

QY 348 tcgcggcgtgtgtgtgattgtt 368
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 314 TCGGCGCTGTGCTGCTCTTG 294

RESULT 12
US-09-103-840A-1/c
: Sequence 1, Application US/09103840A
: Patent No. 6294328
: GENERAL INFORMATION:
: APPLICANT: FLEISCHMAN, Robert D.
: APPLICANT: WHITE, Owen R.
: APPLICANT: FRASER, Claire M.
: APPLICANT: VENTER, John C.
: TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
: FILE REFERENCE: 24366-20007.00

```


OM of: US-09-484-577a-4 to: GenEmbl.* out_format : pfs
 Date: Jul 16, 2002 2:10 AM

About: Results were produced by the Gencore software, version 4.5,
 Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=firmer+2.pn.model -DEV=xlh
 -O=4cgr2.1/USPTO.spool/US09484577/runat.15072002.153752.8137/app-query.fasta_1.184
 -DB=GenEmbl -GPM=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
 -MIMARCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blowm62 -TRANS=human40.cdi
 -LIST=45 -DOCALL=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEADSIZE=500
 -MINLEN=0 -MAXLEN=200000000 -USER=US09484577@CGN1.1_3836
 -NCPU=6 -ICPU=3 -LONGLOG -DEV=TIMEOUT=120 -WARN_TIMEOUT=30
 -NO_XLPHY -WAIT -THREADS=1

Search information block:

Query: US-09-484-577a-4
 Query length: 124
 Database: GenEmbl.*
 Database sequences: 1797656
 Database length: 187333701
 Search time (sec): 1788.330000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation
gb_Ba:AF414151	+	243.00	370.16	3.0e-12	334050	Yersinia pestis str
gb_Ba:AE004049	+	241.00	394.27	1.3e-13	15552	AE004049 Yersinia pestis str
gb_Ba:AF314525	+	238.00	410.48	1.7e-14	1428	AF314525 Yersinia pestis str
gb_Ba:AF121772	+	238.00	400.63	5.9e-13	1428	AF121772 Yersinia pestis str
gb_Ba:AE002524	-	238.00	391.76	1.1e-13	11381	AE002524 Yersinia pestis str
gb_Pat:AX044033	-	238.00	360.85	9.8e-12	349980	AX044033 Sequence 112 from F
gb_Pat:AX044034	-	238.00	360.85	9.8e-12	349980	AX044034 Sequence 113 from F
gb_Ba:NM0622491	-	235.00	356.15	1.8e-11	326301	NM0622491 Yersinia pestis str
gb_Ba:AF646077	-	218.00	330.84	4.6e-10	189050	AF646077 Ralstonia solanaceae
gb_Ba:SSU51197	-	215.00	342.47	1.0e-10	28804	SSU51197 Sphingomonas 588 sp
gb_Pat:AR068625	+	205.00	343.44	8.9e-11	3526	AR068625 Rhizobium leguminosar
gb_Ba:AF314503	+	182.00	296.61	3.7e-08	6952	AF314503 Yersinia pestis str
gb_Ba:AF314506	+	182.00	296.61	3.7e-08	6952	AF314506 Yersinia pestis str
gb_Ba:AF314508	+	182.00	296.61	3.7e-08	6952	AF314508 Yersinia pestis str
gb_Ba:AF314510	+	182.00	296.61	3.7e-08	6952	AF314510 Yersinia pestis str
gb_Ba:AF314512	+	182.00	296.61	3.7e-08	6952	AF314512 Yersinia pestis str
gb_Ba:AF314514	+	182.00	296.61	3.7e-08	6952	AF314514 Yersinia pestis str
gb_Ba:AF314515	+	182.00	296.61	3.7e-08	6952	AF314515 Yersinia pestis str
gb_Ba:AF314517	+	182.00	296.61	3.7e-08	6952	AF314517 Yersinia pestis str
gb_Ba:AF314519	+	182.00	296.61	3.7e-08	6952	AF314519 Yersinia pestis str
gb_Ba:AF314521	+	182.00	296.61	3.7e-08	6952	AF314521 Yersinia pestis str
gb_Ba:AF314523	+	182.00	296.61	3.7e-08	6952	AF314523 Yersinia pestis str
gb_Ba:AF314524	+	182.00	296.61	3.7e-08	6952	AF314524 Yersinia pestis str
gb_Ba:AF314526	+	182.00	296.61	3.7e-08	6952	AF314526 Yersinia pestis str
gb_Ba:AF314528	+	182.00	296.61	3.7e-08	6952	AF314528 Yersinia pestis str
gb_Ba:AF314530	+	182.00	296.61	3.7e-08	6952	AF314530 Yersinia pestis str
gb_Ba:AF314532	+	182.00	296.61	3.7e-08	6952	AF314532 Yersinia pestis str
gb_Ba:AF314534	+	182.00	296.61	3.7e-08	6952	AF314534 Yersinia pestis str
gb_Ba:AF314536	+	182.00	296.61	3.7e-08	6952	AF314536 Yersinia pestis str
gb_Ba:AF314538	+	182.00	296.61	3.7e-08	6952	AF314538 Yersinia pestis str
gb_Ba:AF314540	+	182.00	296.61	3.7e-08	6952	AF314540 Yersinia pestis str
gb_Ba:AF314542	+	182.00	296.61	3.7e-08	6952	AF314542 Yersinia pestis str
gb_Ba:AF314544	+	182.00	296.61	3.7e-08	6952	AF314544 Yersinia pestis str
gb_Ba:AF314546	+	182.00	296.61	3.7e-08	6952	AF314546 Yersinia pestis str
gb_Ba:AF314548	+	182.00	296.61	3.7e-08	6952	AF314548 Yersinia pestis str
gb_Ba:AF314550	+	182.00	296.61	3.7e-08	6952	AF314550 Yersinia pestis str
gb_Ba:AF314552	+	182.00	296.61	3.7e-08	6952	AF314552 Yersinia pestis str
gb_Ba:AF314554	+	182.00	296.61	3.7e-08	6952	AF314554 Yersinia pestis str
gb_Ba:AF314556	+	182.00	296.61	3.7e-08	6952	AF314556 Yersinia pestis str
gb_Ba:AF314558	+	182.00	296.61	3.7e-08	6952	AF314558 Yersinia pestis str
gb_Ba:AF314560	+	182.00	296.61	3.7e-08	6952	AF314560 Yersinia pestis str
gb_Ba:AF314562	+	182.00	296.61	3.7e-08	6952	AF314562 Yersinia pestis str
gb_Ba:AF314564	+	182.00	296.61	3.7e-08	6952	AF314564 Yersinia pestis str
gb_Ba:AF314566	+	182.00	296.61	3.7e-08	6952	AF314566 Yersinia pestis str
gb_Ba:AF314568	+	182.00	296.61	3.7e-08	6952	AF314568 Yersinia pestis str
gb_Ba:AF314570	+	182.00	296.61	3.7e-08	6952	AF314570 Yersinia pestis str
gb_Ba:AF314572	+	182.00	296.61	3.7e-08	6952	AF314572 Yersinia pestis str
gb_Ba:AF314574	+	182.00	296.61	3.7e-08	6952	AF314574 Yersinia pestis str
gb_Ba:AF314576	+	182.00	296.61	3.7e-08	6952	AF314576 Yersinia pestis str
gb_Ba:AF314578	+	182.00	296.61	3.7e-08	6952	AF314578 Yersinia pestis str
gb_Ba:AF314580	+	182.00	296.61	3.7e-08	6952	AF314580 Yersinia pestis str
gb_Ba:AF314582	+	182.00	296.61	3.7e-08	6952	AF314582 Yersinia pestis str
gb_Ba:AF314584	+	182.00	296.61	3.7e-08	6952	AF314584 Yersinia pestis str
gb_Ba:AF314586	+	182.00	296.61	3.7e-08	6952	AF314586 Yersinia pestis str
gb_Ba:AF314588	+	182.00	296.61	3.7e-08	6952	AF314588 Yersinia pestis str
gb_Ba:AF314590	+	182.00	296.61	3.7e-08	6952	AF314590 Yersinia pestis str
gb_Ba:AF314592	+	182.00	296.61	3.7e-08	6952	AF314592 Yersinia pestis str
gb_Ba:AF314594	+	182.00	296.61	3.7e-08	6952	AF314594 Yersinia pestis str
gb_Ba:AF314596	+	182.00	296.61	3.7e-08	6952	AF314596 Yersinia pestis str
gb_Ba:AF314598	+	182.00	296.61	3.7e-08	6952	AF314598 Yersinia pestis str
gb_Ba:AF314600	+	182.00	296.61	3.7e-08	6952	AF314600 Yersinia pestis str
gb_Ba:AF314602	+	182.00	296.61	3.7e-08	6952	AF314602 Yersinia pestis str
gb_Ba:AF314604	+	182.00	296.61	3.7e-08	6952	AF314604 Yersinia pestis str
gb_Ba:AF314606	+	182.00	296.61	3.7e-08	6952	AF314606 Yersinia pestis str
gb_Ba:AF314608	+	182.00	296.61	3.7e-08	6952	AF314608 Yersinia pestis str
gb_Ba:AF314610	+	182.00	296.61	3.7e-08	6952	AF314610 Yersinia pestis str
gb_Ba:AF314612	+	182.00	296.61	3.7e-08	6952	AF314612 Yersinia pestis str
gb_Ba:AF314614	+	182.00	296.61	3.7e-08	6952	AF314614 Yersinia pestis str
gb_Ba:AF314616	+	182.00	296.61	3.7e-08	6952	AF314616 Yersinia pestis str
gb_Ba:AF314618	+	182.00	296.61	3.7e-08	6952	AF314618 Yersinia pestis str
gb_Ba:AF314620	+	182.00	296.61	3.7e-08	6952	AF314620 Yersinia pestis str
gb_Ba:AF314622	+	182.00	296.61	3.7e-08	6952	AF314622 Yersinia pestis str
gb_Ba:AF314624	+	182.00	296.61	3.7e-08	6952	AF314624 Yersinia pestis str
gb_Ba:AF314626	+	182.00	296.61	3.7e-08	6952	AF314626 Yersinia pestis str
gb_Ba:AF314628	+	182.00	296.61	3.7e-08	6952	AF314628 Yersinia pestis str
gb_Ba:AF314630	+	182.00	296.61	3.7e-08	6952	AF314630 Yersinia pestis str
gb_Ba:AF314632	+	182.00	296.61	3.7e-08	6952	AF314632 Yersinia pestis str
gb_Ba:AF314634	+	182.00	296.61	3.7e-08	6952	AF314634 Yersinia pestis str
gb_Ba:AF314636	+	182.00	296.61	3.7e-08	6952	AF314636 Yersinia pestis str
gb_Ba:AF314638	+	182.00	296.61	3.7e-08	6952	AF314638 Yersinia pestis str
gb_Ba:AF314640	+	182.00	296.61	3.7e-08	6952	AF314640 Yersinia pestis str
gb_Ba:AF314642	+	182.00	296.61	3.7e-08	6952	AF314642 Yersinia pestis str
gb_Ba:AF314644	+	182.00	296.61	3.7e-08	6952	AF314644 Yersinia pestis str
gb_Ba:AF314646	+	182.00	296.61	3.7e-08	6952	AF314646 Yersinia pestis str
gb_Ba:AF314648	+	182.00	296.61	3.7e-08	6952	AF314648 Yersinia pestis str
gb_Ba:AF314650	+	182.00	296.61	3.7e-08	6952	AF314650 Yersinia pestis str
gb_Ba:AF314652	+	182.00	296.61	3.7e-08	6952	AF314652 Yersinia pestis str
gb_Ba:AF314654	+	182.00	296.61	3.7e-08	6952	AF314654 Yersinia pestis str
gb_Ba:AF314656	+	182.00	296.61	3.7e-08	6952	AF314656 Yersinia pestis str
gb_Ba:AF314658	+	182.00	296.61	3.7e-08	6952	AF314658 Yersinia pestis str
gb_Ba:AF314660	+	182.00	296.61	3.7e-08	6952	AF314660 Yersinia pestis str
gb_Ba:AF314662	+	182.00	296.61	3.7e-08	6952	AF314662 Yersinia pestis str
gb_Ba:AF314664	+	182.00	296.61	3.7e-08	6952	AF314664 Yersinia pestis str
gb_Ba:AF314666	+	182.00	296.61	3.7e-08	6952	AF314666 Yersinia pestis str
gb_Ba:AF314668	+	182.00	296.61	3.7e-08	6952	AF314668 Yersinia pestis str
gb_Ba:AF314670	+	182.00	296.61	3.7e-08	6952	AF314670 Yersinia pestis str
gb_Ba:AF314672	+	182.00	296.61	3.7e-08	6952	AF314672 Yersinia pestis str
gb_Ba:AF314674	+	182.00	296.61	3.7e-08	6952	AF314674 Yersinia pestis str
gb_Ba:AF314676	+	182.00	296.61	3.7e-08	6952	AF314676 Yersinia pestis str
gb_Ba:AF314678	+	182.00	296.61	3.7e-08	6952	AF314678 Yersinia pestis str
gb_Ba:AF314680	+	182.00	296.61	3.7e-08	6952	AF314680 Yersinia pestis str
gb_Ba:AF314682	+	182.00	296.61	3.7e-08	6952	AF314682 Yersinia pestis str
gb_Ba:AF314684	+	182.00	296.61	3.7e-08	6952	AF314684 Yersinia pestis str
gb_Ba:AF314686	+	182.00	296.61	3.7e-08	6952	AF314686 Yersinia pestis str
gb_Ba:AF314688	+	182.00	296.61	3.7e-08	6952	AF314688 Yersinia pestis str
gb_Ba:AF314690	+	182.00	296.61	3.7e-08	6952	AF314690 Yersinia pestis str
gb_Ba:AF314692	+	182.00	296.61	3.7e-08	6952	AF314692 Yersinia pestis str
gb_Ba:AF314694	+	182.00	296.61	3.7e-08	6952	AF314694 Yersinia pestis str
gb_Ba:AF314696	+	182.00	296.61	3.7e-08	6952	AF314696 Yersinia pestis str
gb_Ba:AF314698	+	182.00	296.61	3.7e-08	6952	AF314698 Yersinia pestis str
gb_Ba:AF314700	+	182.00	296.61	3.7e-08	6952	AF314700 Yersinia pestis str
gb_Ba:AF314702	+	182.00	296.61	3.7e-08	6952	AF314702 Yersinia pestis str
gb_Ba:AF314704	+	182.00	296.61	3.7e-08	6952	AF314704 Yersinia pestis str
gb_Ba:AF314706	+	182.00	296.61	3.7e-08	6952	AF314706 Yersinia pestis str
gb_Ba:AF314708	+	182.00	296.61	3.7e-08	6952	AF314708 Yersinia pestis str
gb_Ba:AF314710	+	182.00	296.61	3.7e-08	6952	AF314710 Yersinia pestis str
gb_Ba:AF314712	+	182.00	296.61	3.7e-08	6952	AF314712 Yersinia pestis str
gb_Ba:AF314714	+					

protein 20020 TR:AGS4319 (EMBL:AB005178) (207 aa) fasta
scores: E(): 6.6, 26.3% id in 156 aa. The similarity is
limited to the central and C-terminal regions of this CDS"

gene
CDS
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAC90808.1"
/db_xref="GI:15980009"
/translation="MAVAVDKENTLITLNEFVSFLETSGTIEKLRNLSYSSEVMSCV
AQLALANKAHIFSIILIELEKTDHQSQNYQAOGEVLYTTPWTIRLMIPGCMIA
HOASLIPSYKRAHDHAFDMLVETFGFGYTSLEGEFNEPIORTGILQLSYKGMEM
LYKRENEVIYFKNADVHIOHEPESLSVSLNLIINKPDASARQVFDLESDEGSNLCM
GKPHNENKTIKASQTSFSLSTGHGMSKRAIKRIAHNHAEVRALSVYAMLRAAD
TVLSVLTLEKSHYRVRVLAELNNAQGETSNAKP
complement(2167..3072)
/gene="YP01996"
complement(2167..3072)
/note="No significant database hits"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAC90809.1"
/db_xref="GI:15980010"
/translation="MSKVTSPDSIRIKRHTVIAASPNSESTISLSYKKEASELLE
KENAFQFDLPLESLKDDIILEINDLEKDTTSNQILSEHVVILSRHNPFLHLM
RLIGSPSTLYLANFEDILVNPNTPISELYOCQVEDISEPSPSLIRIEDHNPILM
GUYEFVEAFYLTDEDEAODAFILILHSPKMLSWYKNSLEPSTICTSLAS
RLQWVRLMGANNAQVAVPELAKTSYANFVRMEAVESLTLDPOSTELDLTVK
NNSDAQITAEASQSLAKAESEV"
complement(3863..4936)
/gene="YP01997"
complement(3863..4936)
/gene="YP01997"
/note="No significant database matches"
/codon_start=1
/transl_table=11
/product="putative exported protein"
/protein_id="CAC90810.1"
/db_xref="GI:15980011"
/translation="MKALLKTFITCMVLGGLISQSTAEPTKPLVIOGSGFAAGAV
ITTPQOPKPKRIDPTQGYHGDHSAVYQIPEPHKKYPIYMLGAQSSSTWMSSTD
GREGONIFLRKRGSTYLYVDOPRKNAGRSVEGTVKFLADEQLFQFRLGPDFF
DSVQVSHKETLNQYYRQMTPTGTFPFDIVLSDANSAVDSGPAILLFTSGGGPCWM
FLTKMNNANVKGIVAREPSSVEYFEGVDVPPPTKNAFDVVKGESVPLEKFMALITPIV
IYGNIDIDPKPTALPAQDSMRARLIVARQMRDVYVNHGQDVTYIHLPEIGIKGNTHP
FSDLNVEELADVSKEFLKDNLQ"
complement(5149..5619)
/gene="YP01998"
complement(5149..5619)
/gene="YP01998"
complement(5149..5619)
/note="Similar to Escherichia coli Peps1 TnpA TR:P75026
(EMBL:U60777) (140 aa) fasta scores: E(): 0.13, 31.4% id
in 137 aa, and to Thermotoga maritima hypothetical 15.1
kDa protein Tm1010 TR:Q340A3 (EMBL:AB001162) (135 aa)
fasta scores: E(): 1.2e-12, 42.0% id in 119 aa"
/codon_start=1
/transl_table=11
/product="putative exported protein"
/protein_id="CAC90811.1"
/db_xref="GI:15980012"
/translation="MKVFLITLALLLPQTQALAAQAAYTVSPAGSQPTTIGAESEF
TGTAVDSRFRGTAPARISGTVSEAGAPRAPHMTHPLGOTLIYSSGTGWQMEGAA
QOIKIGDVAIVPQYKGVHMGASAKESVMHIAVSSLDQNTVTWMEKVTDDGYP"
complement(5674..6096)
/gene="YP01999"
complement(5674..6096)
/gene="YP01999"
complement(5674..6096)
/EC_number="4.1.1.44"
/note="Similar to Acinetobacter calcoaceticus
4-carboxymuconolactone decarboxylase Pcad SW:DD4C.ACTIA
(P20370) (134 aa) fasta scores: E(): 1.4e-09, 35.9% id in
103 aa, and to Methanobacterium thermoautotrophicum

```

gamma-carboxymuconolactone decarboxylase Mch234 TR:O263333
(EMBL:AE000810) (123 aa) fasta scores: E(): 4.7e-11, 35.5%
id in 107 aa"
/codon_start=1
/transl_table=11
/product="putative decarboxylase"
/protein_id="CAC90812.1"
/db_xref="GI:15960013"
/transl_table="MMKK1QQYLCTALTAPGASMASIAHAODAKNNYKOPSPAQDL
MGDIAPMAQLTIDVLADWIDIERPOLSKRDRSLVTSALVAMDRPDRLSHLRARON
GVPEELVEIATOLCAFSGMPTAVTAISAKEYKEKG"
complement(6373 .7764)
/gene="YP02000"
complement(6373 .7764)
/gene="YP02000"
complement(6373 .7764)
/EC_number="2.7.3.-"
/note="Similar to Burkholderia pseudomallei sensor protein
Irls SW:IRLS.BURPS (O31396) (464 aa) fasta scores: E(): 0,
29.1% id in 454 aa, and to Pseudomonas aeruginosa probable
two-component sensor PAL438 TR:AG604827 (EMBL:AE004573)
(481 aa) fasta scores: E(): 0, 42.2% id in 453 aa"
/codon_start=1
/transl_table=11
/product="putative two-component system sensor protein"
/protein_id="CAC90813.1"
/db_xref="GI:15960014"
/transl_table="MMIKRISGRILAMFALASILVSLGIVLRSSLNHSLOKMH
NELOFTESIMAPWILSRITLDDQMOTLAKFTDLAAAGSGVQVSLDDPOLQVGPFA
PRGVTAISKSGKSPGKVPAGNDESCILFSLSELTPAERAPALRYVVAIDSTPYMGNTLN
EFTFTLVITATIGVLPFALLGFIYATRGIMRPVKNLSQAOHLARNGHNGQDLITLPLD
ELQGLAFAENGVLEROETLAWKQLESFNADVAHNEIRPLTLNLQCTQLGTSKRTHFL
EDLLESNLEELERMTSTVNDMLFLSHAQGEHNAHQLTVDVSLRETYKTYTVSEFAKE
KNLSIHVEGDITATIDRRFLFRSLANLLENSARVAHGVTVKVTISNNMLASIAVAN
IGAPIAEHLNRLREPERYRVDSNRASDTHHGLSLIVRAVALMHRDGVASSQNGIN
TFGLTATDPYI"
complement(6334 .7017)
/gene="YP02000"
/note="Pfam match to entry PF00512 signal, Histidine
kinase, score 129.50, E-value 6.1e-35"
complement(7057 .7269)
/gene="YP02000"
/note="Pfam match to entry PF00672 DUF5, HAMP domain,
score 52.70, E-value 7.8e-12"
complement(7213 .7278)
/gene="YP02000"
/note="1 probable transmembrane helices predicted for
YP02000 by TMHMM2.0"
complement(7770.8462)
/gene="copr"
/note="copr"
/note="YP02001"
complement(7770.8462)
/gene="copr"
/note="copr"
/note="Similar to Pseudomonas syringae transcriptional
activator protein CPR SW:COPR.PSESM (002540) (227 aa)
fasta scores: E(): 0, 60.8% id in 227 aa, and to
Burkholderia pseudomallei transcriptional activator
protein Irlr SW:IRLR.BURPS (O31395) (229 aa) fasta scores:
E(): 0, 57.6% id in 224 aa"
/codon_start=1
/transl_table=11
/product="transcriptional activator"
/protein_id="CAC90814.1"
/db_xref="GI:15960015"

```



```

CDS
3867. .4112
/ gene="XF2399"
/ note="hypothetical protein; identified by sequence
similarity; putative; ORF located using Glimmer/RBSfinder"
/ codon_start=1
/ transl_table=11
/ product="hypothetical protein"
/ protein_id="AAE85198.1"
/ db_xref="GI:9107582"
/ translation="MPINTTASHKTHRPCCOHCNDASATATPTTPTPTPTL
KDSRASPIAHLTACLRPPPTKHHTAHVPIYITK"
4428. .5057
/ gene="XF2400"
4428. .5057
/ gene="XF2400"
/ note="hypothetical protein; identified by sequence
similarity; putative; ORF located using Glimmer/RBSfinder"
/ codon_start=1
/ transl_table=11
/ product="hypothetical protein"
/ protein_id="AAE85199.1"
/ db_xref="GI:9107583"
/ translation="MTMTANVFKKILKICGMVIALMAFGAFISGIIYIEKAOQ
DATAQIKESNETILQSDSSYKRCPEKDELATGLGIIIFNNELYSVEK
SPSKPHGPSSVTIGIFGKYSKKNYINSYNQDNKQITVIGKQIELLYEFSNT
YSAMDNDLFIYIKADGQPKNKKRIHELKILIDLEAHITPCOGES"
5054. .5326
/ gene="XF2401"
5054. .5326
/ gene="XF2401"
/ note="hypothetical protein; identified by sequence
similarity; putative; ORF located using Glimmer/RBSfinder"
/ codon_start=1
/ transl_table=11
/ product="hypothetical protein"
/ protein_id="AAE85200.1"
/ db_xref="GI:9107584"
/ translation="NMTSSYFEDATLDRAEYRSGNCPGMYQILSDAVKAGGDA
RLANGKLNAGDHSACGRTPIRAMDQADREKNINTQHHPAL"
5329. .5946
/ gene="XF2402"
5329. .5946
/ gene="XF2402"
/ note="hypothetical protein; identified by sequence
similarity; putative; ORF located using Glimmer/RBSfinder"
/ codon_start=1
/ transl_table=11
/ product="hypothetical protein"
/ protein_id="AAE85201.1"
/ db_xref="GI:9107585"
/ translation="MNSPRVFKSLKLGILATJLVMEFGAILSCGNMTEETTPORA
AAODIKESKMMLOSRNSKRCPEPPEWVNOFSPKAIFFKNIYLSIETSP
GNLGEAGTSIVSLIPDKSKKEYIHARLPGSNQOVEKGFYFYNNLDFLXS
AIDKGLFNEYEYAGDGLPRNKKERIRELKIIDFVAHITPCOGES"
5943. .6215
/ gene="XF2403"
5943. .6215
/ gene="XF2403"
/ note="hypothetical protein; identified by sequence
similarity; putative; ORF located using Glimmer/RBSfinder"
/ codon_start=1
/ transl_table=11
/ product="hypothetical protein"
/ protein_id="AAE85202.1"
/ db_xref="GI:9107586"
/ translation="MNTSLDFYDQATLERANYNNGDMNYQTLSAIVNAGGDA
RLANGFNAGDHSFQRDTPSIRAMDADREKNINTQHHPAL"
6218. .6841
/ gene="XF2404"
6218. .6841
/ gene="XF2404"
/ note="hypothetical protein; identified by sequence
similarity; putative; ORF located using Glimmer/RBSfinder"

```

```

/ codon_start=1
/ transl_table=11
/ product="hypothetical protein"
/ protein_id="AAE85203.1"
/ db_xref="GI:9107587"
/ translation="MNSPRVFKSLKLGILATJLVMEFGAILSCGNMTEETTPORA
AAODIKESKMMLOSRNSKRCPEPPEWVNOFSPKAIFFKNIYLSIETSP
GNLGEAGTSIVSLIPDKSKKEYIHARLPGSNQOVEKGFYFYNNLDFLXS
AIDKGLFNEYEYAGDGLPRNKKERIRELKIIDFVAHITPCOGES"
6838. .7110
/ gene="XF2405"
6838. .7110
/ gene="XF2405"
alignment_scores:
  Quality: 241.00      Length: 98
  Ratio: 3.347        Gaps: 0
  Percent Similarity: 73.469      Percent Identity: 53.061
alignment_block:
US-09-484-577A-4 x AE004049
Align seg 1/1 to: AE004049 from: 1 to: 15552
20  LysThrArgAspGluLeuAlaPheLeuProAlaLeuGluIleValG1 36
    ::::| | | | | | | | | | | | | | | | | | | | | | | |
2495 CGCAGCGCGATGAGCGACCTTCCTACCGCCCATGAGAACTGACCGA 2544
    36 uThrProSerProThrAlaArgLeuThrAlaLeuLeuAlaAla 53
    | | | | | | | | | | | | | | | | | | | | | | | |
2545 AACCCCTGTCCTCCCTACCGCGAGCATGCGCATATTATTTGCT 2594
    53 eupPtyrCysAlaValAlaPheValGlyLeuGlyArgIleAspIleVal 69
    | | | | | | | | | | | | | | | | | | | | | | | |
2555 GCTTATGCTGCGCTCTGCTGCGCTGCTGCGGAAATTAATTTGTT 2644
    70 AlaSerAlaSerArgLysIleValProGlyAspArgValLysLeuValG1 86
    | | | | | | | | | | | | | | | | | | | | | | | |
2645 GCGGTGCGCGCGGGAATGCTGCTGACACGACGACCAAGTACTCCA 2694
    86 nProLeuGluValGlyValAlaArgAlaThrHisValArgAspGlyIleT 103
    | | | | | | | | | | | | | | | | | | | | | | | |
2695 ACCACTAGACGCGCGGTAGTACGCGCTATCTGATCGCGAGTGGCGAA 2744
    103 hrValLysAlaGlyIleLeuIleGluLeuAspProPheAla 117
    | | | | | | | | | | | | | | | | | | | | | | | |
2745 AAGTGAAGCACACCAACTGATGCAACTGATGCCACCGCC 2788
seq_name: gb_ba:NM391263
seq_documentation_block:
LOCUS NM391263 1428 bp DNA linear BCT 29-MAR-2000
DEFINITION Neisseria meningitidis hlyd gene (putative component of type I
secretion system), strain FAM18.
ACCESSION AJ391263 GI:6900444
VERSION AJ391263.1
KEYWORDS hlyd gene.
SOURCE Neisseria meningitidis.
ORGANISM Neisseria meningitidis.
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
REFERENCE 1 (bases 1 to 1428)
AUTHORS Klee,S.R., Nassif,X., Kusecek,B., Merker,P., Beretti,J.L.,
Achtman,M. and Tinsley,C.R.
TITLE Molecular and biological analysis of eight genetic islands that
distinguish neisseria meningitidis from the closely related
pathogen neisseria gonorrhoeae
JOURNAL Infect. Immun. 68 (4), 2082-2095 (2000)
MEDLINE 20187481
REFERENCE 2 (bases 1 to 1428)
AUTHORS Achtman,M.
TITLE Direct Submision
JOURNAL Submitted (19-JAN-2000) Achtman M., Abt. Trautner,

```



```

source
    3200..4254
    /note="insertion sequence; similar to the putative
    /synchocystis PCC6803 transposase and small repeated
    elements of Neisseria gonorrhoeae"
    /organism="Neisseria meningitidis"
    /db_xref="taxon:487"

BASE COUNT      1070 a      1021 c      1130 g      1032 t      1 others
ORIGIN

alignment_scores:
    Quality:      238.00      Length:      103
    Ratio:        3.216      Gaps:      0
    Percent Similarity: 71.845      Percent Identity: 49.515

alignment_block:
    us-09-484-577a-4 x AF121772 ..

Align seg 1/1 to: AF121772 from: 1 to: 4254

20  LysThrArgAspGluLeuAlaPheLeuProAlaAlaLeuIleValG1 36
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
311 CGCAGCGCGGAGAAACAGCGCTTTTCCCGCGCATTTGGAACTGACCGA 360
    36 uThrProSerProThrAlaArgLeuThrAlaAlaLeuAlaAlaL 53
    ::::: ::::: ::::: ::::: ::::: ::::: :::::
361 TACGGCGGCTCTGCGCTCCGAAATGGCGCGGCTTTATTATGCGCT 410
    53 euPheTyrCysAlaValAlaTrpAlaGlyLeuGlyArgGleAspIleVal 69
    ::::: ::::: ::::: ::::: ::::: ::::: :::::
411 TTGGCGCTTTGGCTTGTGTTGCTGCTGCTGCGCAAAATCATATGTG 460
    70 AlaSerAlaSerArgLysIleValProGlyAspArgValLysLeuValG1 86
    ::::: ::::: ::::: ::::: ::::: ::::: :::::
461 GCGGCGGCTGCGGCGCAAAACGCTGCGGCGCGCGCAGCAAAACCATCCA 510
    86 nProLeuGluAlaGlyValAlaArgAlaThrHisValArgAspGlyGlnT 103
    ::::: ::::: ::::: ::::: ::::: ::::: :::::
511 GCGCTGGAAACGGCGGTGTAAAGCGGTACATGTGCGCGCGGCGCAGC 560
    103 hrValLysAlaGlyIleLeuIleGluLeuAspProPheAlaGlyGly 119
    ::::: ::::: ::::: ::::: ::::: :::::
561 ATGTCAACAGCAGGAAACGCTGCGCGCAACTGAGAGCTGTGGACACAGAC 610
    120 ValAspVal 122
    :::::
611 AGCGATGTG 619

name: gb_ba:AE002524

seq documentation block:
LOCUS      AE002524      11381 bp      DNA      linear      BCT 25-MAY-2000
DEFINITION Neisseria meningitidis serogroup B strain MC58 section 166 of 206
ACCESSION  AE002524 AE002098
VERSION    AE002524.1 GI:7226991
KEYWORDS
SOURCE
ORGANISM   Neisseria meningitidis MC58.
            Neisseria meningitidis MC58
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE  1 (bases 1 to 11381)
AUTHORS   Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C.,
            Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F.,
            Dodson,R.J., Nelson,W.C., Gwinn,M.L., DeBoy,R., Peterson,J.D.,
            Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D.,
            Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E.,
            Clitane,H., Clark,E.B., Cotton,M.D., Uterback,T.R., Khouri,H.,
            Qin,H., Yamathavan,J., Gill,J., Scarlato,V., Masignani,V.,
            Pizzo,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R.,
            Rappuoli,R. and Venter,J.C.
            Complete genome sequence of Neisseria meningitidis serogroup B
            strain MC58

JOURNAL    Science 287 (5459), 1809-1815 (2000)
MEDLINE    2017575
PUBMED     10710307
REFERENCE  2 (bases 1 to 11381)
AUTHORS   Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C.,
            Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F.,
            Dodson,R.J., Nelson,W.C., Gwinn,M.L., DeBoy,R., Peterson,J.D.,
            Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D.,
            Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E.,
            Clitane,H., Clark,E.B., Cotton,M.D., Uterback,T.R., Khouri,H.,
            Qin,H., Yamathavan,J., Gill,J., Scarlato,V., Masignani,V.,
            Pizzo,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R.,
            Rappuoli,R. and Venter,J.C.
            Direct Submission
            Submitted (17-MAR-2000) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
            Location/Qualifiers
                1..11381
                /organism="Neisseria meningitidis MC58"
                /strain="MC58"
                /db_xref="taxon:122586"
                /note="serogroup: B"
                165..976
                /gene="NMB1736"
                /note="This region contains an authentic point mutation,
                causing a premature stop, and is not the result of a
                sequencing artifact. This region contains an authentic
                frame shift and is not the result of a sequencing
                artifact.; similar to PID:1653908 PID:1651692"
                complement(1140..2543)
                /gene="NMB1737"
                complement(1140..2543)
                /gene="NMB1737"
                /note="similar to PID:1139569 percent identity: 48.26;
                identified by sequence similarity; putative"
                /transl_table=1
                /product="secretion protein, putative"
                /protein_id="AAF42081.1"
                /db_xref="GI:7226992"
                /translation="MTLLNLMIMQDYGISVCTLTPTLYQLHFLFSAMKSYFSKYILPVS
                LFTPLSLSPSVSAFTLPEAMRAAQHSDPFOASHQORAVRARRQDAKAFIPHYSA
                NASYQAPSPSISSTRETQSGVVGQTLFDPAFAQYRQSFQAAEQAEAFAREEL
                LKVAESYFNVLSDPTVAHAHAKEAYAOQVQAQALFPKGAATLADTHEAAGRON
                ALAQLAVLAERQYENQNDYTDLSKQTEADLIDTARLYPKTERYSIDWQBITA
                LSNNHEYRMQQLALOSSGQALRAQNSRPTVAHNGYQNNLTSSAQNNDIYRKKG
                MSYVOLNLPITYGGEISGKIHAEQYGAERQQLTATERHITLAVRQAYTESGAR
                YQIAQERVLSSRLKLTSTGTGOYGINRNLEVIARQEAQAOKLAQARKFMILAY
                LRLVKSGLGLTFVFAE"
                complement(2612..4039)
                /gene="NMB1738"
                complement(2612..4039)
                /gene="NMB1738"
                /note="similar to GP:558154 percent identity: 65.47;
                identified by sequence similarity; putative"
                /codon_start=1
                /transl_table=11
                /product="secretion protein, putative"
                /protein_id="AAF42082.1"
                /db_xref="GI:7226993"
                /translation="MFESALKSFSLRYITVWRNVAVRDQLKPPKRTAEQAFPLAHL
                ELTDPVSAAPKWAARFIMAFALALMLSMFQKIDIVAASGKTVSGGRSKTQPLET
                AVKAVHVDGQHVKGERTLALEAVLQSDVYQSOALQAOALSLREAVIALAES
                RIVPHIDMAQARSISLADVQSAOVLQHOVQANVAQAOALQSLRGEQALQSKAKA
                GQKPLVSGAITEQKTRADVRRLRANRFSSEHAFLEQOQSVSNMNDLESTROMRQIO
                AALIAEQNRVNTQNKRTDIDALRQAEQIDQYRGQTDKAKQROQMLTQSPADGT
                VQELATYTGCVQAQAQKMVTAPDDDKMDVVLVLMKDITGFEVQGDVAVVETSPY
                TRYGUTGKVSQSDAVSHQGLGVYAVVSLDKHTLINDKAVNLTAGMVAETAEIK
                TGRRLVDTLPSLPQTKLDESFRER"
                4275..4628
                /gene="NMB1739"
                4275..4628
                /gene="NMB1739"

```

[illegible]

seq_documentation_block:

LOCUS AX044033 349980 bp DNA linear PAT 24-NOV-2000

DEFINITION Sequence 112 from Patent WO0066791.

ACCESSION AX044033

VERSION AX044033.1 GI:11342917

KEYWORDS

SOURCE

ORGANISM

Neisseria meningitidis.

Neisseria meningitidis

Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;

Neisseria.

REFERENCE

1 (bases 1 to 349980)

Plaza, M., Hickey, E., Peterson, J., Tettelin, H., Venter, J. C., Maignani, V., Galeotti, C., Mora, M., Ratti, G., Scariselli, M., Scarlato, V., Rappuoli, R., Frazer, C. M. and Grandi, G.

Neisseria genomic sequences and methods of their use

Patent: WO 0066791-A 112 09-NOV-2000;

JOURNAL

CHIRON CORPORATION (US); THE INSTITUTE FOR GENOMIC RESEARCH (US)

FEATURES

Source

1. 349980

Location/Qualifiers

1. 349980

/organism="Neisseria meningitidis"

/db_xref="taxon:487"

/note="sequence too long, cut in 8 pieces.-seq 1: 1 to 349980 349980 bases-seq 108: 300001 to 649980 349980 bases-seq 109: 600001 to 949980 349980 bases-seq 110: 900001 to 1249980 349980 bases-seq 111: 1200001 to 1549980 349980 bases-seq 112: 1500001 to 1849980 349980 bases-seq 113: 1800001 to 2149980 349980 bases-seq 114: 2100001 to 2272325 172325 bases"

BASE COUNT 87189 a 93501 c 84627 g 84663 t

ORIGIN

alignment_scores:

Quality: 238.00 Length: 103

Ratio: 3.216 Gaps: 0

Percent Similarity: 71.845 Percent Identity: 49.515

alignment_block:

US-09-484-577A-4 x AX044033/rev ..

Align seg 1/1 to reverse of: AX044033 from: 1 to: 349980

20 LysThrArgAspGluLeuAlaPheLeuProAlaAlaLeuGluIleValG1 36

CGCACGGGGGAGAAACAGCGCTTTTGGCCCGCATTTGGAACGACCGA 323780

36 uThrProSerProThrAlaArgLeuThrAlaAlaLeuAlaAlaL 53

TACGCCGCTCTCGCCGCTCGCAAAATGGCGCGCTTTATATATGCGCT 323730

53 eupHeTyrcysAlaValaIatrpAlaGlyLeuGlyArgIleAspIleVal 69

TTGGCGCTTTGGCTTTGTTGGTCCTGTCGCGCAAAATCGATATGTG 323680

323729 TTGGCGCTTTGGCTTTGTTGGTCCTGTCGCGCAAAATCGATATGTG 323680

70 AlaSerAlaSerArgLysIleValProGlyAspArgValLysLeuValG1 86

CGCGCGCTTTGGCGCAAAACGCTGCGCGCGCGCGCACCAAAACCATCCA 323630

323679 GCGCGCGCTTTGGCGCAAAACGCTGCGCGCGCGCGCACCAAAACCATCCA 323630

86 nProLeuGluValGlyValaIatrpAlaThrHisValaArgAspGlyG1nt 103

GCCCGTGAACAGCGCGGTGTTAAGCGGTACATGTGCGCGACGCGCAC 323580

323629 GCCCGTGAACAGCGCGGTGTTAAGCGGTACATGTGCGCGACGCGCAC 323580

103 hrValLysAlaGlyLleuIleuIleGluLeuAspProPheAlaGlyGly 119

ATGTGAAACAGGAGAAACGCTGCGCAAACTGAGAGCTGTGGAAACAGAC 323530

323579 ATGTGAAACAGGAGAAACGCTGCGCAAACTGAGAGCTGTGGAAACAGAC 323530

120 ValAspVal 122

AGCGATGTG 323521

323529 AGCGATGTG 323521

seq_name: gb_pat:AX044034

seq_documentation_block:

LOCUS AX044034 349980 bp DNA linear PAT 24-NOV-2000

DEFINITION Sequence 113 from Patent WO0066791.

ACCESSION AX044034

VERSION AX044034.1 GI:11342918

KEYWORDS

SOURCE

ORGANISM

Neisseria meningitidis.

Neisseria meningitidis

Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;

Neisseria.

REFERENCE

1 (bases 1 to 349980)

Plaza, M., Hickey, E., Peterson, J., Tettelin, H., Venter, J. C., Maignani, V., Galeotti, C., Mora, M., Ratti, G., Scariselli, M., Scarlato, V., Rappuoli, R., Frazer, C. M. and Grandi, G.

Neisseria genomic sequences and methods of their use

Patent: WO 0066791-A 113 09-NOV-2000;

JOURNAL

CHIRON CORPORATION (US); THE INSTITUTE FOR GENOMIC RESEARCH (US)

FEATURES

Source

1. 349980

Location/Qualifiers

1. 349980

/organism="Neisseria meningitidis"

/db_xref="taxon:487"

/note="sequence too long, cut in 8 pieces.-seq 1: 1 to 349980 349980 bases-seq 108: 300001 to 649980 349980 bases-seq 109: 600001 to 949980 349980 bases-seq 110: 900001 to 1249980 349980 bases-seq 111: 1200001 to 1549980 349980 bases-seq 112: 1500001 to 1849980 349980 bases-seq 113: 1800001 to 2149980 349980 bases-seq 114: 2100001 to 2272325 172325 bases"

BASE COUNT 86473 a 95646 c 85908 g 81953 t

ORIGIN

alignment_scores:

Quality: 238.00 Length: 103

Ratio: 3.216 Gaps: 0

Percent Similarity: 71.845 Percent Identity: 49.515

alignment_block:

US-09-484-577A-4 x AX044034/rev ..

Align seg 1/1 to reverse of: AX044034 from: 1 to: 349980

20 LysThrArgAspGluLeuAlaPheLeuProAlaAlaLeuGluIleValG1 36

CGCACGGGGGAGAAACAGCGCTTTTGGCCCGCATTTGGAACGACCGA 23780

36 uThrProSerProThrAlaArgLeuThrAlaAlaLeuAlaAlaL 53

TACGCCGCTCTCGCCGCTCGCAAAATGGCGCGCTTTATATATGCGCT 23730

53 eupHeTyrcysAlaValaIatrpAlaGlyLeuGlyArgIleAspIleVal 69

TTGGCGCTTTGGCTTTGTTGGTCCTGTCGCGCAAAATCGATATGTG 23680

23729 TTGGCGCTTTGGCTTTGTTGGTCCTGTCGCGCAAAATCGATATGTG 23680

70 AlaSerAlaSerArgLysIleValProGlyAspArgValLysLeuValG1 86

CGCGCGCTTTGGCGCAAAACGCTGCGCGCGCGCGCACCAAAACCATCCA 23630

23679 GCGCGCGCTTTGGCGCAAAACGCTGCGCGCGCGCGCACCAAAACCATCCA 23630

86 nProLeuGluValGlyValaIatrpAlaThrHisValaArgAspGlyG1nt 103

GCCCGTGAACAGCGCGGTGTTAAGCGGTACATGTGCGCGACGCGCAC 23580

23629 GCCCGTGAACAGCGCGGTGTTAAGCGGTACATGTGCGCGACGCGCAC 23580

103 hrValLysAlaGlyLleuIleuIleGluLeuAspProPheAlaGlyGly 119

ATGTGAAACAGGAGAAACGCTGCGCAAACTGAGAGCTGTGGAAACAGAC 23530

23579 ATGTGAAACAGGAGAAACGCTGCGCAAACTGAGAGCTGTGGAAACAGAC 23530

120 ValAspVal 122

AGCGATGTG 23521

23529 AGCGATGTG 23521

seq_name: gb_ba:NME391260

seq_documentation_block:

LOCUS NME391260 7824 bp DNA linear BCT 29-MAR-2000

DEFINITION Neisseria meningitidis DNA for region 7 (homology to type I secretion system) and flanking genes, strain Z2491.

ACCESSION AJ391260.1 GI:6900414

VERSION 1

KEYWORDS GTP-pyrophosphokinase; hlyD gene; ORF-C7; outer membrane protein; relA gene; rseI gene; tolC gene; transposase.

SOURCE Neisseria meningitidis.

ORGANISM Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.

REFERENCE 1 (bases 1 to 7824)
Klee, S.R., Nassif, X., Kusecek, B., Merker, P., Beretti, J.L., Achtman, M., and Tinsley, C.R.
Molecular and biological analysis of eight genetic islands that distinguish neisseria meningitidis from the closely related pathogen neisseria gonorrhoeae
Infect. Immun. 68 (4), 2082-2095 (2000)

JOURNAL 2 (bases 1 to 7824)
Achtman, M.
Direct Submission
Submitted (19-JAN-2000) Achtman M., Abt. Trautner, Max-Planck-Institut fuer molekulare Genetik, Ihmestr. 73, Berlin, 14195, GERMANY

FEATURES
Location/Qualifiers
1..7824
/organism="Neisseria meningitidis"
/strain="Z2491"
/db_xref="taxon:487"
/country="Gambia"
/note="serogroup A, subgroup IV-1"
/note="ORF-C7; homology to AB015669; designated NM1999 in the Sanger Center Z2491 genome sequence"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB72024.1"
/db_xref="GI:6900415"
/translation="MLAEVFGFGITKESGSGRFRYRCWLMGSEVLYGRVHFGCO
NNTILFELTGCGVAKGEMESRLFAFLNIAIPKTRVDIADFEENGYSPPNARD
RNGMETHCHHVRKCEGLSDMEDEDAKMTKGTGIGRSRSKYRVYKRGOLD
KSTWTREKFEKADIVIPFVLONPGEYFGAYPICERFVQKATRIHAKEDKVIS
ADRYLEWFKKGRANGLKPIFELDKALFELIEPSHKLAPLAPAYDCAFLLA
QAIHERPAKPKYKDPYMYEYENLEKQLEQKHVNNEESYNNFTYKFAPIISMA"
1456..6701
/note="Region 7, a DNA island present in N. meningitidis and absent from N. gonorrhoeae"
2510..3874
/gene="hlyD (left fragment)"
2510..3874
/note="hlyD (left fragment)"
/note="homology to hemolysin secretion protein D P09986; 98% identical to NatD component of the type I secretion system AD30977; designated natD (NM1996) in the Sanger Center genome sequence"
/codon_start=1
/pseudo
/transl_table=11
/note="insertion/duplication that inactivates hlyD gene"
3907..4071
/gene="hlyD (right fragment)"
3907..4071
/note="hlyD (right fragment)"
/note="designated natD (NM1995) in the Sanger Center Z2491 genome sequence"
/codon_start=1
/pseudo
/transl_table=11
/gene="tolC"
4140..5543
CDS

/gene="tolC"
/note="homology to tolC, P02930; 99% identical to NatC, component of the type I secretion system AD30978; designated natC (NM1994) in the Sanger Center Z2491 genome sequence"
/codon_start=1
/transl_table=11
/product="putative outer membrane protein"
/protein_id="CAB72027.1"
/db_xref="GI:6900416"
/translation="WTLLNLMIMODYGISVCLTTPYIQHLEFSAMKSYFSYIIPVS
LFTPLSLSPSSVSAFTLPEANRAAOHSADQASHYORDAVARAQQAALPHVSA
NAYROPPSISSTRETQGSVQVOTLFDQAKFAQYQSRDTPQAEGRPDAEEL
LTKVAESYENVLLSRDITVAANAEEKAAVOOQALFNKGAATLIDHEKAGYDN
ALAOEINVLAKORYENOLINDYTGDSMOIEAIDPNTLARYLPKLEYSNDYEMORIA
LSNNHETMOQLAOSQALRAANSRPTYSAHVGYONNTYTSOANDYHYGKG
MSVGQVLNPLYTGELSGKTHEAEQAQGAAPLATERIKLAVROQYTSGAR
QIMAOERVLSSRLKSTETGQOYGINRLEVIAROEVAQAEOKLAQARYKFLAY
LRLVKSGLGLETFPAE"
complement(5707..6042)
/gene="rseI"
complement(5707..6042)
/note="homology to putative transposase of CAB44496; designated NM1992 in the Sanger Center Z2491 genome sequence"
/codon_start=1
/transl_table=11
/product="putative transposase"
/protein_id="CAB72028.1"
/db_xref="GI:6900417"
/translation="MSAOGNRLIAPMYQNTMGVFEAFMFQOCLPALTKSVIIL
DNARPRMGVLRMAEKIGKVLPLAPYLPENLPEKVMANIKRYLRYTSLDVARDD
ALLSTFDPN"
6514..6619
/note="C106, Correia element"
complement(6787..7824)
/gene="relA"
complement(6787..7824)
/gene="relA"
/function="synthesis of ppGpp, a mediator of the stringent response"
/note="homology to GTP-pyrophosphokinase of P11585; designated relA (NM1991) in the Sanger Center Z2491 genome sequence"
/codon_start=1
/transl_table=11
/product="GTP-pyrophosphokinase"
/protein_id="CAB72029.1"
/db_xref="GI:6900418"
/translation="MAESGKEDLAFAKTELENDITYVLTLPHGKVLSPGATPIFA
VALHSSIGDRGARGVEGIVLSPLENGOREITLAKGHSVMALEKREDDYTTAYCO
IGKTRATYIQONADYVREGGRVQLDKQLKLPKPMLOLAELNLSYKREDDYTTAYCO
GEISNRATOKACGTLNPPVSEYTIYKOSITKGGKGVLLDEGDLMTTLACG
KPAEDYDVGVTNRGIVSHKRTCSFPHLHAPEKVLDAEWALQEGVFAVDIE
IRAQSRGILRDVSDALAHKLNVTAVQOSRDLASMRFTLEVKQVTDLPVLTASIG
DVKGVLSTVRL"

BASE COUNT 1853 a 1854 c 2065 g 2052 t
ORIGIN

alignment_scores:
Quality: 235.00 Length: 103
Ratio: 3.219 Gaps: 0
Percent Similarity: 70.874 Percent Identity: 49.515

alignment_block:
US-09-484-577A-4 x NME391260 ..
Align seg 1/1 to: NME391260 from: 1 to: 7824

20 LysThrArgAspGluLeuAlaPheLeuProAlaAlaLeuGluIleValG1 36
:::||||| |||||||:::.....

[illegible]

```

      /db_xref="SPRREMBL:O9JUK7"
      /translab="M1NTAKATFEGRNRLIKOVKRSVARINALBEOQALSDADLO
      AKTAEKORLABDSQTIDGILPEAPACREASRPRLMRHPDVLIGGMVLIHQKIEM
      RPTGGSTLVATLAIVTLNALAGKVHYTVYNDYLASHDAGIMPELYNPFLITGVIIISD
      MOFEDRONAAYADITYGTNNFEGFDLRDMWVDODTKVORELNFRNVDEYBSIIDE
      ARPLILISQGDADNIOLYOIMNTVPPLLHMLHAARHLTFEKDHVAIIOGEIIVDFE
      ABOILITOMGLIANDSLYSANALINLMALRAHTLFEPKDYHAI1SAGHEH
      TGLMSGRMSEGIHOAVEAKEEGEVETFOYFELRYTRLSGMTGTADTE
      AFPEOSIYNLEVIPIFNRPVKORDENDOLFRAEEKEFELVVADIIECHKRQGPIVG
      TTIENSELSVSHLQKAGLPHNVLNKKEHREALIYAQAQKGVALIVATMMAGRGTDI

alignment_scores:
    Quality: 235.00          Length: 103
    Ratio:   3.219           Gaps: 0
    Percent Similarity: 70.874     Percent Identity: 49.515

alignment_block:
US-09-484-577A-4 x NMA6Z2491/rev ..

Align seg 1/1 to reverse of: NMA6Z2491 from: 1 to: 326301

20 LysThrArgAspGlulLeuAlaPheLeuProAlaIleLeuValgl 36
:::|||||::::||| ||||| ||||| ||||| :::::
273552 CCACGCCGCGGAAGAACAAGCCGTTTGCCCCGCATTGGAACTGACCCA 273503
36 uThrProSerProThraIArgLeuThraIalaLeuAlaIal 53
||||| ||| ::::: ::::: ||| |::|
273502 TAGCGCGGTCGTGCCGCTCCGAAATGGCGCGCGCTTTATTATGGCGT 273453
53 euPheTrcysAlaValaIatrpAlaClyLeucylArgIleaspIleval 69
::: ||::: ||::: ::::: ||::: ||::: ||||| ||||| |||||
273452 TTGCGGCTTTGGCTTGTGTGGTCCGTGCTGGCAAAATCGATATTGTG 273403
70 AlaSerAlaSerArgIlysIleValProGlyAspArgValIysLeuValgl 86
|||:::||||| ||| ||| ||| ||| ||| ||| ::::|
273402 GGCGGCGCTTCGGGCAAAAACGGTGTGGGGGCGCACGCAAAACCATCCA 273353
86 nProLeuGluValgIyValaIarqAlahrhIsValaIarqAspGlylgt 103
||||| ::::: ||||| ::::: ||||| ::::: ||||| :::::
273352 GGCGCTGGAAACGCTGTGGTTAAGCGCGTACACTTGCGGAGCGGCGACG 273303
103 hrValIysAlaGlyGluIleLeuIleGluLeuAspProPheAlaClygly 119
||||| ||||| ||| ||||| :::::
273302 ATGTGAACAGGAGAACGCTGTGGCAACTGGAGCTGTGGGAACAGAC 273253
120 ValAspVal 122
|||||
273252 AGCGACTGTG 273244

seq_name: gb_ba:AL646077
seq_documentation_block:
LOCUS AL646077 189050 bp DNA linear BCT 07-DEC-2001
DEFINITION Ralstonia solanacearum GM11000 megaplasmid, complete sequence;
ACCESSION AL646077 AL646053
VERSION segment 2/11.
KEYWORDS AL646077.1 GI:17430642
SOURCE Ralstonia solanacearum.
ORGANISM Ralstonia solanacearum
Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
Ralstonia.
REFERENCE 1 (bases 1 to 189050)
Salanoubat,M., Genin,S., Artiguenave,F., Gouzy,J., Mangenot,S.,
Arlat,M., Billaut,A., Broffier,P., Camus,J.C., Catolico,L.,
Chandler,M., Choisine,N., Claudel-Renaud,C., Cunac,S., Demange,N.,
Gaspin,C., Lavie,M., Moisan,A., Robert,C., Sautin W.S., Schlex,T.,
Siglier,P., Thebaull,P., Whalen,M., Wincker,P., Levy,M.,
Weissenbach,J. and Boucher,C.A.
Genome sequence of the plant pathogen Ralstonia solanacearum
Unpublished
2 (bases 1 to 189050)
Boucher,C.A.
```


predicted by Framed"

/codon_start=1
/transl_table=1
/evidence=not-experimental
/product="HYPOTHETICAL PROTEIN"
/protein_id="CAD17330.1"
/db_xref="GI:17430647"
/translation="MEANADPRKQPPMRVGRCAATATLSALAPALADAPKPGPMO
DEPARAVSEAOPEVSGTHPLVLIHADPAMVAQVSTDFVHKNVAIPKGRSLI
GKELRKVNAREIWIQIIPASGTLRLDPLQATMRDSAGVADLSPGALLGAMTS
EPFVPH"

gene
complement(8066..8488)
/gene="RSP0180"
/note="RSD0692"
complement(8066..8488)
/gene="RSP0180"
/function="miscellaneous; hypothetical/global homology"
/note="Product confidence : hypothetical
Gene name confidence : hypothetical
Predicted by Framed"
/codon_start=1
/transl_table=1
/evidence=not-experimental
/product="CONSERVED HYPOTHETICAL PROTEIN"
/protein_id="CAD17331.1"
/db_xref="GI:17430648"
/translation="MGAWSHASFGNDADIDFAEVAETNDISRVESAFDAVLAAGDD
LEAPEASQVAEVAEVARLQHWGVRNADSDISAMIERGTITPRDIHAKRQALDR
IVCEPSELLEIOMSEDDKRDVARELRARLDGAPAR"
complement(8552..8956)
/gene="RSP0181"
/note="RSD0693"

gene

alignment_scores:
Quality: 218.00 Length: 113
Ratio: 2.868 Gaps: 1
Percent Similarity: 67.257 Percent Identity: 41.593

alignment_block:

US-09-484-577A-4 x AL646077/rev ..

Align seg 1/1 to reverse of: AL646077 from: 1 to: 189050

12 ArgGluValValAlaValAlaGlyLysThrArgAspLeuLeu..... 26
|||||
163572 CGGGATGTCCTCCGCCACTACTGGAACGACGCGCGCTGAGCCTGCC 163523
27PheLeuProAlaAlaLeuGluLeu 35
163522 CCGCTTCGAGGTGACAGCGCGGATTCCTGCCCGCGCGCTGCGCTGC 163473
35 AlGluThrProProSerProThrAlaArgLeuThrAlaAlaLeuLeuAla 51
163472 AGGCGCACCGCGTCTCGCGCGCGGCGGATGGGTCGCGCGCTGATG 163423
52 AlaLeuThrCysAlaValAlaAlaTrrAlaGlyLeuGlyArgIleAsp 68
163422 CTGCTGATTCGGTGCCTGCTGCTGCTGCTGCTGCGCAAGTGGACAT 163373
68 eValAlaSerAlaSerArgLysIleValProGlyAspArgValLysLeu 85
163372 CATGTCACGCGGCGGCAAGATCATCCGTCGACGCGGCAAGACCA 163323
85 AlGlnProLeuGluValGlyValAlaArgAlaThrHisValArgAspGly 101
163322 TTGCGCGCGGTGGAGTGGCCACCGTGGCGCAAGCTGACGTGAGGAGGA 163273
102 GlnThrValLysAlaGlyGluIleLeuIleGluLeuAsp 114
163272 CAGATCGTCAGAGCGCGGCGGATGCTGATCGACGCTGAC 163224
seq_name: gb_ba:SS051197

seq documentation block:

LOCUS SS051197 28804 bp DNA linear BCT 16-MAY-1996
DEFINITION Sphingomonas S88 sphingan polysaccharide synthesis (spsG), (spsS),
(spsR), glycosyl transferase (spsQ), (spsI), glycosyl transferase
(spsK), glycosyl transferase (spsL), (spsJ), (spsF), (spsD),
(spsC), (spsE), Urf 32, Urf 26, ATP-binding cassette transporter
(atrb), ATP-binding cassette transporter (atrb),
glucosyl-1-isoprenylphosphate transferase (spsB), glucose-1-phosphate
thymidyl transferase (rhaS), dTDP-6-deoxy-D-glucose -3,5-epimerase
(rhaC) dTDP-D-glucose 4,6-dehydratase (rhaB),
dTDP-6-deoxy-L-mannose-dehydrogenase (rhaD), Urf 31, and Urf 34
genes, complete cds.
ACCESSION U51197
VERSION U51197.1 GI:1314561
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLES
JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLES
JOURNAL
FEATURES
source
1. 28804
/organism="Sphingomonas sp. S88"
/strain="S88"
/db_xref="taxon:46624"
complement(325..1944)
/gene="spsG"
complement(325..1944)
/gene="spsC"
complement(325..1944)
/function="sphingan polysaccharide synthesis"
/codon_start=1
/transl_table=1
/evidence=experimental
/product="unknown"
/protein_id="AAC44055.1"
/db_xref="GI:1314562"
/translation="MATSARVSGSGGLAVVOPIRARDIVFPARVILYHPQSG
LVTLGRFGLAVIVFVSVYGLSLILPSPMLGAAFLALIALFVIALPEADKAP
TKLLRCYVLFMILVMPNYGLSLVAGLPWISFRLEFVLTLLVCSISPFR
DMAVTLASPMILAREMLGFLIIQIILISLISQSPSISGLVINSILVFPFATMLCQ
GNSRTPKEMVVRNMLGCVGLVILVIGFEEFRAEHLVWGIIPSLQINDPSVILHLTSC
VGTGYVNVNPGNPLVNGELVMTIPFALHLMVNSVSAALALITLFDLAVMLVSAWS
GARLSVNGVAVAHATYLLILGLRLRMTKREGGLVYSTMMVYPAFAVVLVIMVAV
HNRVLGGATVACSDNGRHQIRKGEFLKLMRPVFGCPQSSAEVYVGRSQSGTSLSDS
GFLSTAVDVGVLGFLAFYGTWVGAQLLRAGLTDSQGPLHLAVATTVVLLNTRL
VLSQINDVGVFMVLGFLFALYRSKQARGA"
complement(1955..3313)
/gene="spsS"
complement(1955..3313)
/gene="spsS"
complement(1955..3313)
/function="sphingan polysaccharide synthesis"
/codon_start=1
/transl_table=1
/evidence=experimental
/product="unknown"
/protein_id="AAC44056.1"
/db_xref="GI:1314563"
/translation="MRSGLPPEGPCPKKAKQTSPPVGOAFNVTRSKHLKLRQWVG
FVKGAGAVTSAFTWLVARAFGPAVGAFGLTTRQMLVILSLGLDTILVTRVAV
CYSQNRGVARTIRAHARTVAVAGGLVLMVMIILFQOQIATNLSGPPAPADLLNGF
LIPVSACRLASTTTRMGMDIGKSNWDGLTGLTGLGAMLAGALLIGVANHPLPSVL

```

gene
CDS
    complement(313..5325)
    /gene="spsr"
    /complement(313..5325)
    /gene="spsr"
    /function="sphingan polysaccharide synthesis"
    /codon_start=1
    /transl_table=11
    /product="unknown"
    /protein_id="AAC44057.1"
    /db_xref="GI:1314564"
    /translation="MPDILVRNSELDAIRKAGGERTKIAAGYTSTYVTTFTT
    PYTISLASKRPVNTTLOIINSSNLYKYNVAQFOPADQYNTANRTIGSSNVVD
    GYTLGGTGDTYKSKMGILLRESKGVITKNSVDHFGALNNVLTVDGVTLLQNNFHD
    NRRDHTNREASNVVINDNMFTNLYPVNHEPDAIQFTNGTKGNTITITNNYMG
    GDGLAFQGVFMGNESIPEYNTIKNNLYQNGFTGHNVLGKNVYDSNTIISMD
    ALADWIRLDGVTGKVTNNVDOVVITATSGITQSNNAVLATDAVTLRKIGDLNAGS
    KQALNLIYSGVQOPAGSAFAQVAKQAAQSTGANILLDLOFTATGIVDQTRM
    STDFTKALDLTAISNGFQVCTGTGFEIIRNSQIRLPAFTLNFDMKRTATGAV
    GOIMGVNOSWGISLADGELVYTVYNAAGQSTVATSGAKMTDTAHKIALITYDSAGK
    KAITVDVAVKSGTIVGSTRAVEYAGLYIGSPNAVSGSVGEIEMRDALASQIIL
    ALMAATASOPTTIVKPSLTSLSAASVTLIGTPTVTALPAATGTLTIGGSASTA
    SVAQSLAQAISLMAAGTAASATGSLSVANSLPATKLVQDFYHV"
    5526..6473
    /gene="spsq"
    5526..6473
    /gene="spsq"
    /function="sphingan polysaccharide synthesis"
    /codon_start=1
    /transl_table=11
    /evidence="experimental"
    /product="glycosyl transferase"
    /protein_id="AAC44058.1"
    /db_xref="GI:1314565"
    /translation="MEASPTDVSILVAVYASAPITGICRIGIAAAGTAEHILLID
    NNGGTEAVVAEFPFHVRIPESENGIGAGNRCARARAPRLILVNDALPRGAI
    DLVAFKAHPDAAGRSYEPNGOLDHNPILPTVRDFVSISSSPMRGSLPA
    DATAGPEVINGGPMVYDARVMTIDSGDFGLXSEIDLFCQIRKRGISLVDP
    AGVVDHGGGHSISPTRYLPLITGGMHARHFGHVGAVTGMALMAAAKVVYVIGL
    LGRISPRRAAMNLRDMSITVFGQPRKMHGWRDVT"
    complement(6558..7460)
    /gene="spsr"
    complement(6558..7460)
    /gene="spsr"
    /function="sphingan polysaccharide synthesis"
    /codon_start=1
    /transl_table=11
    /product="unknown"
    /protein_id="AAC44059.1"
    /db_xref="GI:1314566"
    /translation="MKKLIYTVAAVAAALVAGSGSKGKIDKGVAVYVSDGETTF
    ELNAEVOATOVPOSTDRKLAQMALQRIETKIKTARBEKDKTPEVTLIQRRAD
    LIITTLRLDKLAGIAQPTDAEIAQTEAHADPRQKRTISIEQIVPPSSAKFE
    FAPLTKLDQLAKLTAGVOFRAPQDLTALPPEAKIAAPAEFLILPQOGV
    TANYITGTTIOLAGDQARELALNTLTERFSKADQLNRLKKARESVYQGYSP
    PLOGGAGAAAPAAPAPMAAQ"
    7588..8646
    /gene="spsk"
    7588..8646
    /gene="spsk"
    /function="sphingan polysaccharide synthesis"
    /codon_start=1
    /transl_table=11
    /evidence="experimental"
    /product="glycosyl transferase"
    /protein_id="AAC44060.1"
    /db_xref="GI:1314567"
    /translation="MGDMAEATVTEAKGPKLMCLASGGHLROILDESVMKEH
    DYFVEDTALGRSLAKSHSVALVDHVALQAKIGHRLNLGGMWRRLRSLSIIRKH

```

```

gene
CDS
    complement(9552..10940)
    /gene="spsr"
    complement(9552..10940)
    /gene="spsr"
    /function="secretion of sphingan polysaccharide"
    /codon_start=1
    /transl_table=11
    /evidence="experimental"
    /product="unknown"
    /protein_id="AAC44062.1"
    /db_xref="GI:1314569"
    /translation="WAVGSAKFLMPFGRDVPEDGYEPLAATVPPRDAHAGRGR
    DPEFPRASALDRPLDRRERREITRAREFALATFTPPQPADRSSFAGRLVFLRLI
    SAIESORSHVLYGEGIGKTSILHLVDVARESSYISATCGANANPVLIDFAYLE
    DVLPEHFGSVNAGEAGSGSLADRLTPGDSDELADICATGCTPVLLIDFDR
    VTDPRFOOVAELIKNISDRSARQOLYAGVSNQOELVGYPSIRRNITGIPMRLE
    ETEVQEMALGETASGLFDPRLRITILHLGSPYFARLCHNALDALDQGRITVD
    LGLHRLADQALILETGRMSPRAVIEMKFEVGRYDPLVALGASRSTDMFTQOAV
    VELLGSLTAAVOEQLDALSRQLESDADQDERRRFNDLSIPVYLMVGRIRI
    DSGLEELAAY"
    11569..12867
    /gene="spsr"
    11569..12867
    /gene="spsr"
    /function="sphingan polysaccharide synthesis"
    /codon_start=1
    /transl_table=11
    /evidence="experimental"
    /product="unknown"
    /protein_id="AAC44063.1"
    /db_xref="GI:1314570"
    /translation="MKPEGPGEVYVAVKDVQOAVERYLGRDRVSRCRYLALATATAIQ
    PALQORAFTRPGRSGSERQISVHATQLEVDNVNVLNDRITTSARGGVIASPILDS
    IVPRAFGOLYLAGTVGYREYRRTYNNRENIISLGGDQRIASCYVHGEYGRHLT

```

alignment_scores:

Percent Similarity:	Ratio:	Length:	Gaps:	Percent Identity:
76.289	2.905	97	0	45.361

alignment_block:

US-09-484-577A-4 x SSU51197/rev ..

Align seg 1/1 to reverse of: SSU51197 from: 1 to: 28804

```

27 pheleuProAlaLeuGluIleValGluThrProProSerProThrAl 43
|||||
18808 TTCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 18759
43 aargleuThrAlaLeuLeuAlaLeuLeuAlaLeuLeuAlaLeu 60
|||||

```

```

18758 GCGCCTACCGCGCGGTGATGTCGGCGGCTGGCATCCACCGCCT 18709
60 rpAlaGlyLeuGlyArgIleAspIleValAlaSerAlaSerArgLysIle 76
18708 GGTGGCAATCGCCGCGTGAAGTGTGGCGCCAGCGCCGCGCATC 18659
77 ValProGlyAspArgValLysLeuValGlnProLeuGlnValGlyValVa 93
18658 GCGCCGATCGCGAGACCAAGATCGTCGATCCCGGAAGCGGTATCGT 18609
93 lArGAlaThrHisValArgAspGlyGlnThrValLysAlaGlyGluIle 110
18608 CCGGCCATTCGTGTGGCGCGAGGCGAGGTTTCAGAAAGCGCAGGTGC 18559
110 euIleGluLeuAspProPheAlaGlyLysValAla 123
seq_name: gp_pat:AR068625

documentation block:
DEFINITION Sequence 1 from patent US 5854034.
ACCESSION AR068625
VERSION AR068625.1 GI:6000832
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 28804)
AUTHORS Pollock,T.J., Yamazaki,M., Thorne,L., Mikolajczak,M. and Armentrout,R.W.
TITLE DNA segments and methods for increasing polysaccharide production
JOURNAL Patent: US 5854034-A 1 29-DEC-1998;
FEATURES
location/Qualifiers
source 1..28804
BASE COUNT 4974 a 9804 c 9230 g 4796 t
ORIGIN

alignment_scores:
Quality: 215.00 Length: 97
Ratio: 2.905 Gaps: 0
Percent Similarity: 76.289 Percent Identity: 45.361

alignment_block:
US-09-484-577A-4 x AR068625/rev ..
lgn seg 1/1 to reverse of: AR068625 from: 1 to: 28804

27 PheLeuProAlaAlaLeuGluIleValGlnThrProProSerProThrAl 43
18808 TTCTTCGCGCGCGCGCTCGAGATCATCGAACGTCGCGTGCACCGC 18759
43 aArgLeuThrAlaAlaLeuLeuAlaLeuPheTyCysAlaValAlar 60
18758 GCGCCTACCGCGCGCGGTGATGTCGGCGGCTGGCATCCACCGCCT 18709
60 rpAlaGlyLeuGlyArgIleAspIleValAlaSerAlaSerArgLysIle 76
18708 GGTGGCAATCGCGCGCGTGAAGTGTGGCGCCAGCGCGCATC 18659
77 ValProGlyAspArgValLysLeuValGlnProLeuGlnValGlyValVa 93
18608 CCGGCCATTCGTGTGGCGCGAGGCGAGGTTTCAGAAAGCGCAGGTGC 18559
93 lArGAlaThrHisValArgAspGlyGlnThrValLysAlaGlyGluIle 110
18608 CCGGCCATTCGTGTGGCGCGAGGCGAGGTTTCAGAAAGCGCAGGTGC 18559
110 euIleGluLeuAspProPheAlaGlyLysValAla 123

```

```

18558 TGATCAAGCTCGATCCACCATGTGCGCGACCGAGCGCGG 18518
seq_name: gp_ba:AF273216
seq_documentation_block:
LOCUS AF273216 3526 bp DNA linear BCT 17-JUL-2001
DEFINITION Rhizobium leguminosarum bv. viciae bacteriocin (rzca) gene, partial
cds: ABC transporter Rzcb (rzcb) gene, complete cds; and
transporter component Rzcd (rzcd) gene, partial cds.
ACCESSION AF273216
VERSION AF273216.1 GI:10952793
KEYWORDS
SOURCE
ORGANISM Rhizobium leguminosarum bv. viciae.
Rhizobium leguminosarum bv. viciae
Bacteria: Proteobacteria; alpha subdivision; Rhizobiaceae group;
Rhizobiaceae; Rhizobium.
REFERENCE 1 (bases 1 to 3526)
AUTHORS Venter,A.P., Twelker,S., Oresnik,I.J. and Hynes,M.F.
TITLE Analysis of the genetic region encoding a novel rhizobiocin from Rhizobium leguminosarum bv. viciae strain 306
JOURNAL Canadian journal of microbiology. 47 (6), 495-502 (2001)
MEDLINE 21360352
PUBMED 11467725
REFERENCE 2 (bases 1 to 3526)
AUTHORS Venter,A.P., Twelker,S. and Hynes,M.F.
TITLE Direct Submission
JOURNAL Submitted (30-MAY-2000) Biological Sciences, University of Calgary, 2500 University Drive NW, Calgary, AB T2N 1N4, Canada
FEATURES
source
location/Qualifiers
1..3526
/organism="Rhizobium leguminosarum bv. viciae"
/strain="305"
/db_xref="taxon:387"
/clone="pRle306c"
/note="extrachromosomal
biovat: viciae"
<1..76
/gene="rzca"
<1..76
/codon_start=2
/transl_table=11
/product="ABC transporter Rzcb"
/protein_id="AAG25076.1"
/db_xref="GI:10952794"
/translation="MMDIOTASAGVAEVEIRADSPPTPATIADSGVAICAVAGYR
IASRPTLSRELATAPARADLIRAKTIGKARVRAEKISRLATIPAPALASR
GTFAYFAGIAAGSRRLINPIDFSANVADDELATSEPIIVQRRFPGASQONF
GRFWELPAIMRYRRAFGHVLASLVQITALVYPLFQVAVKVALAHSTYLLIVLV
GLAAVGLFDVQVQLFVALSHSTNRIDIVELORRLRLNLPLSFETRAGQIVAR
IRELETRNMLTGOGLFSGLDLFTIIEFVLSYSKLAIVASIPFYVAIGFLIR
PLKERIDKFERGASQOLLEVTVVGIOTLKASAVEPVSSQMEERLAAYRSFLA
TWLAKGONAIQVYVNTTSAILLFGAQVVDIGESGALVAFNMTAGOVSPILRS
QIMODPOVQVSIARISDLNAPQERPSVANSLPPKAIAPKSNFYSPPSGDYL
KQINFIRGEVIGLVPSGSKSTLKLQVRYLPNGQVFDGDIAQVDPAMLR
NIGVAVQENMLFENRTIHNNICMVNPAMSRAAIQMARSGADQFATKLPKRYDTLE
ERGANISGORRLAIARALATNPPLIILIDEATSLDYSERIILGNNMEIYGRFTVI
ITAHRLATYRHSKMDGRIVEGTHETTLARPNGLYAHLMQLOTPIES"
2348..>3526
/gene="rzcd"
2348..>3526
/gene="rzcd"

```

```

/note="required for bacteriocin secretion"
/codon_start=1
/transl_table=11
/product="transporter component Rzcd"
/db_xref="GI:10952796"
/translation="MNQIVQPPRLPPKRLQVNEFLPALEILELPSPRTALI
WFLCFTAGALIMSYIGFDIVATAGKIOPTRGVKALQISIVGRTIAPVSNKAVE
TGDILVEDLPTEARVENVLTLSALRAEVRAALAEVNAWORGDMSTGRRIE
VPEFADSPAEIRREBELLYTSLDLSALSTLDSLAORNOCPAAVKRYTEMTAQAL
VATLADRVAMBSNIYDIDISAGSRGVDAVEITQKEENTLAEIGCARAEKTVTAAS
EELKAIKTFVADNAKQASREIDKEQDLVKAARKRESMTIKSPIINGIVQTSITT
VGQVVTAGELMRIVDASLEIEVLPNRDIGVSPQPAVIVKVEAFPTRYGILNG
T"

BASE COUNT      672 a      1151 c      1025 g      678 t
ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2370
2371
2372
2373
2374
2375
2376
2377
2378
2379
2380
2381
2382
2383
2384
2385
2386
2387
2388
2389
2390
2391
2392
2393
2394
2395
2396
2397
2398
2399
2400
2401
2402
2403
2404
2405
2406
2407
2408
2409
2410
2411
2412
2413
2414
2415
2416
2417
2418
2419
2420
2421
2422
2423
2424
2425
2426
2427
2428
2429
2430
2431
2432
2433
2434
2435
2436
2437
2438
2439
2440
2441
2442
2443
2444
2445
2446
2447
2448
2449
2450
2451
2452
2453
2454
2455
2456
2457
2458
2459
2460
2461
2462
2463
2464
2465
2466
2467
2468
2469
2470
2471
2472
2473
2474
2475
2476
2477
2478
2479
2480
2481
2482
2483
2484
2485
2486
2487
2488
2489
2490
2491
2492
2493
2494
2495
2496
2497
2498
2499
2500
2501
2502
2503
2504
2505
2506
2507
2508
2509
2510
2511
2512
2513
2514
2515
2516
2517
2518
2519
2520
2521
2522
2523
2524
2525
2526
2527
2528
2529
2530
2531
2532
2533
2534
2535
2536
2537
2538
2539
2540
2541
2542
2543
2544
2545
2546
2547
2548
2549
2550
2551
2552
2553
2554
2555
2556
2557
2558
2559
2560
2561
2562
2563
2564
2565
2566
2567
2568
2569
2570
2571
2572
2573
2574
2575
2576
2577
2578
2579
2580
2581
2582
2583
2584
2585
2586
2587
2588
2589
2590
2591
2592
2593
2594
2595
2596
2597
2598
2599
2600
2601
2602
2603
2604
2605
2606
2607
2608

```


gene
DSDGNDKLSFSDSNLKDITFEKYNHNVITTSKKEKVTIONMREADPAKPVNPAT
KDEKTEIIGONGERITSKYDDLIAGNGKRTIODELSKYVDNTELLKHSNANVNSLD
KLISVSATFSNDSRNVLAPISMLDSSLOPARAA"
3399. .5525
/gene="1ktb"
/allele="1ktb1.1"
3399. .5525
/gene="1ktb"
/codon_start=1
/transl_table=1
/product="Lktb"
/protein_id="AAL12767.1"
/db_xref="GI:15987892"

CDS
/translation="MEANHORNDLIVATIMLAOYHNISINPEIKKEDLDGKSLT
TAMLIASLALKAHKEISRLHLVLPALWODNKHLLVYVDNNRYLTNL
BODAPQILSODEFEACYOGLIVTSRVSVOGLAKKFPFTWIPVIRKLETL
VSIQIIFALITPLEFQVMDKLVHRGFSSTNITVALATYIIEIVLGSRTVS
HSTRIDVELAKLERHLLSPISEENRRVDVARELDQINFLTGALTSVLD
LLESPFFAVMWSYPRKLTIVLIGSLPILMSIFISPLRRLEKFAQNDQAF
VESYTAIIMIKAMAYAPQMTDMDKOLASYSRFRVLAITIGOGVOLIOKTMVI
NMLGAHLVIGSDLSIGLIAPNMLSGOVIATVITLAQMODFOOVGISYRELGVIN
SPTEYOQKLSLPEIKGDISEKNIRFRKPPAPITLNNVNEIRGEEVIGYGRSGS
KSTLTKLQREYIPENGQVLDGHDALADPNMLRQIGVLIQDVNLNRSIRENIAL
SDGPMERVIYAAKLAGADHPISELREGYNTVGEQAGLSGGORRITAIRALVN
PKLIFEDNTSALDESEHIIWONOKICOGRTVILLIHLRSTVKNADRIIYMERGEI
VQCKHHEILNSNGLYSTHQLQLN"
5537. .>6952
/gene="1ktb"
/allele="1ktb1.1"
5537. .>6952
/gene="1ktb"
/codon_start=1
/transl_table=1
/product="Lktb"
/protein_id="AAL12768.1"
/db_xref="GI:15987893"

BASE COUNT 2311 a 1219 c 1432 g 1990 t
ORIGIN

Alignment_scores:
Quality: 182.00 Length: 103
Ratio: 2.493 Gaps: 0
Percent Similarity: 70.874 Percent Identity: 35.922

alignment_block:
US-09-484-577A-4 x AF314503 ..

Align seg 1/1 to: AF314503 from: 1 to: 6952

20 LysThrArgAspGluLeuAlaPheLeuProAlaAlaLeuGluIleValG1 36
::: ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
5636 AGAAAAAAGACGAAAGTGAATTTTACCGGACATTTAGAACTGATGGA 5685
36 uThrProSerProThrAlaArgLeuThrAlaLeuLeuAlaAla 53
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
5686 AACCCCGGTTTAAACCAACGCTATATTCCTTATTGATTATGCTAT 5735
53 eupheryCysAlaValAlaTrrpAlaGlyLeuGlyArgIleAspIleVal 69
::: ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
5736 TTTTGGTTGTGCAATGTGTCGACAGTGAAGCAACTGCAATGATG 5785
70 AlaSerAlaSerArgLysIleValProGlyAspArgValLysLeuValG1 86
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

5786 GCGACTGTCGCCGTAATTAATTACTTTAGTGCAGAGTAAGAAATTAA 5835
86 nProLeuGluValGlyValAlaArgAlaThrHisValArgAspGlyG1nt 103
::: ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
5836 ACCGATTGAAAGACGCCATTTGTACAGAAATTTTCGTTAAAGATGGCAGT 5885
103 hrValysAlaGlyGluIleLeuIleGluLeuAspProPheAlaGlyGly 119
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
5886 TTGTGGAAAAAGGCAATTTATGTCAGCTTAAGTGCATGGGTTGAT 5935
120 ValAspVal 122
::: ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
5936 GCAGATATTC 5944

seq_name: gb_ba:AF314506

seq_documentation_block:

LOCUS AF314506 6952 bp DNA linear BCT 21-DEC-2001
DEFINITION Mannheimia haemolytica strain PH8 lktCABD operon, partial sequence.
ACCESSION AF314506
VERSION AF314506.2 GI:15987894
KEYWORDS

SOURCE

ORGANISM

REFERENCE
AUTHORS
TITLE
1 (bases 1 to 6952)
Davies, R.L., Whittam, T.S. and Selander, R.K.
Sequence diversity and molecular evolution of the leukotoxin (lktA)
gene in bovine and ovine strains of Mannheimia (Pasteurella)
haemolytica
J. Bacteriol. 183 (4), 1394-1404 (2001)

REFERENCE
AUTHORS
TITLE
2 (bases 1 to 6952)
Davies, R.L., Campbell, S. and Whittam, T.S.
Mosaic Structure and Molecular Evolution of the Leukotoxin Operon
(lktCABD) in Mannheimia (Pasteurella) haemolytica, Mannheimia
glucosida, and Pasteurella trehalosi
J. Bacteriol. 184 (1), 266-277 (2002)

REFERENCE
AUTHORS
TITLE
3 (bases 1 to 6952)
Davies, R.L., Whittam, T.S. and Selander, R.K.
Direct Submission
Submitted (18-OCT-2000) Institute of Biomedical and Life Sciences,
Division of Infection and Immunity, University of Glasgow, Joseph
Black Building, University Avenue, Glasgow G12 8QU, Scotland
4 (bases 1 to 6952)

REFERENCE
AUTHORS
TITLE
Direct Submission
Submitted (04-OCT-2001) Institute of Biomedical and Life Sciences,
Division of Infection and Immunity, University of Glasgow, Joseph
Black Building, University Avenue, Glasgow G12 8QU, Scotland

REMARK
COMMENT
FEATURES
Sequence update by submitter
On Oct 9, 2001 this sequence version replaced gi:11762023.
Location/Qualifiers
1. .6952
/organism="Mannheimia haemolytica"
/strain="PH8"
/serotype="A1"
/db_xref="taxon:75985"
<1. .>6952
/gene="lktCABD operon"
<1. .447
/gene="lktC"
/allele="lktC1.1"
<1. .447
/gene="lktC"
/codon_start=1
/transl_table=1
/product="LktC"
/protein_id="AAL12769.1"
/db_xref="GI:15987895"

CDS

Tue Jul 16 05:46:30 2002

us-09-484-577a-4.rge


```

seq_name: gb_gss:AF075952
seq_documentation_block:
LOCUS AF075952 750 bp DNA linear GSS 29-AUG-2000
DEFINITION AF075952 Salmonella typhimurium LT2, Lambda DASH II Salmonella
typhimurium genomic clone 1256-r3, DNA sequence.
ACCESSION AF075952
VERSION AF075952.1 GI:3320822
KEYWORDS GSS.
SOURCE Salmonella typhimurium.
ORGANISM Salmonella typhimurium.
Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;
Salmonella.
REFERENCE 1 (bases 1 to 750)
AUTHORS Wong, R.M.-Y., Wong, K.K., Benson, N.R. and McClelland, M.
TITLE Sample sequencing of a Salmonella typhimurium LT2 lambda library:
JOURNAL comparison to the Escherichia coli K12 genome
MEDLINE FEMS Microbiol. Lett. 173 (2), 411-423 (1999)
COMMENT 99243757
Contact: McClelland M
Molecular Biology
Sidney Kimmel Cancer Center
3099 Science Park Road, San Diego, CA 92121, USA
Email: mclelland@iscsi.sdsu.edu
Class: shotgun.
FEATURES
source Location/Qualifiers
1..750
/organism="Salmonella typhimurium"
/db_xref="taxon:602"
/clone_lib="1256-r3"
/clone_lib="Salmonella typhimurium LT2, Lambda DASH II"
/note="Vector: Lambda DASH II; sequenced using Li-Cor
sequencer."
BASE COUNT 160 a 201 c 226 g 163 t
ORIGIN
alignment_scores:
Quality: 105.00 Length: 67
Ratio: 2.143 Gaps: 0
Percent Similarity: 73.134 Percent Identity: 35.821
alignment_block:
US-09-484-577A-4 x AF075952 ..
Align seg 1/1 to: AF075952 from: 1 to: 750
49 LeuLeuAlaAlaLeuPheTyrCysAlaValAlaTPrAlaGlyLeuGlyAr 65
::: |||||
377 ACTGCTCACTGTTATTCTCATTCCTGCGCATCTGCGCTTGGCCG 426
65 gTleAspLleValAlaSerAlaSerArgLysLleValProGlyAspArgV 82
|||||
427 ACTGATGAGGTTCCACCGG.CAACGGAAGATGATCCCACTTCACCGG 475
82 allYsLeuValGlnProLeuGluValAlGlyValAlaValArgAlaThrHisVal 98
|||||
476 AACAGTCTCGACATGCGTGCATGCGGCGCATCTGCGCAGATTGACGGTG 525
99 ArgAspGlyGlnThrValLysAlaGlyLuuLleuLleGluLeuAspPr 115
|||||
526 CGGAGGAGGACAGATTCAGGCTAACCGATGTCGCCGCCGCTGATCC 575
115 o 115
576 G 576
seq_name: gb_est2:BE564254
seq_documentation_block:
LOCUS BE564254 945 bp mRNA linear EST 15-AUG-2000

```

```

DEFINITION 601343055P1 NIH_MGC_53 Homo sapiens cDNA clone IMAGE:3685428 5',
mRNA sequence.
ACCESSION BE564254
VERSION BE564254.1 GI:9807974
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 945)
AUTHORS NIH-MGC http://mhc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps@emall.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Inceye Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov
Plate: LNCM373 row: m column: 13
High quality sequence stop: 5.
FEATURES
source Location/Qualifiers
1..945
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3685428"
/clone_lib="NIH_MGC_53"
/tissue_type="carcinoma, cell line"
/note="Organ: Bladder; Vector: pDNR-LIB (Clontech);
Site_1: Sfil (ggcgcttgccg); Site_2: Sfil (ggcattggcc
5'); Double-stranded cDNA was prepared from cell line RNA.
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-ATTCTAGAGCGCGGAGCGCCGACATG-drr30)BN-3'
sequence: 5'-ATTCTAGAGCGCGGAGCGCCGACATG-drr30)BN-3'
insert size 1.55 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
BASE COUNT 239 a 262 c 354 g 90 t
ORIGIN
alignment_scores:
Quality: 95.50 Length: 120
Ratio: 1.308 Gaps: 6
Percent Similarity: 60.833 Percent Identity: 29.167
alignment_block:
US-09-484-577A-4 x BE564254/rev ..
Align seg 1/1 to reverse of: BE564254 from: 1 to: 945
1 LeuSerSerLeuSerProArgMetLysSerAlaArgGluValAlaVal 17
||| |||||
440 TTGTTCTCTGTCGTCGCCGCTTCCTGCTCTGTTGTCCTGCTG 391
17 lGlyGlyStrArgAspLleuAlaPheLeuProAlaAlaLeuGlu 33
|||||
390 TAGCTCTCTGCGTCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 341
34 lLeValGluThrProProSerProThrAlaArgLeuThrAlaAla..... 48
|||||
340 TTGATT.....CTTCCTCTTCATCTGCTGCTGCTGCTGCTGCTG 297
49 .....LeuLeuAlaAlaLeuPheTyrCysAlaVal 58
|||||
296 CCGGCTGTCGTTGTCGCGCTTCGCTCGACCGGCTCGCTGCTGCTG 247
58 AlAlaTPrAlaGlyLeuGlyArgLleAspLleValAlaSerAlaSerArg 74

```



```

37  ThrProSerProThrAlaArgLeuThrAlaAlaLeuAlaAlaLe 53
    ::::: ||||| ::::: ||||| ::::: |||
21  TCCCGGAGTCCCGCATGAGCAAGAGGCGGCTC.....GCCGG 64
    ::::: ||||| ::::: ||||| ::::: |||
53  uPheTYrCysAlaValAlaTrpAlaGlyLeuGlyArgIleAspIleValA 70
    ||||| ||||| ||||| ||||| |||||
65  CCATGCTGCGCGCTC.....GAGCTGCAG 90
    ::::: ||||| ::::: ||||| ::::: |||
70  lAsSerAlaSerArgIleValProGlyAspArgValIleValGln 86
    ::::: ||||| ::::: ||||| ::::: |||
91  TTGGCGCTACC.....ATTGGCGCGCGCATGCCGTATACCATCGAG 134
    ::::: ||||| ::::: ||||| ::::: |||
87  .....ProLeuGln.....ValGlyValAla 94
    ||||| ||||| ||||| ||||| |||||
135  TCGATGAGAGTGAATCCCGCTCAGTCGAGCCAGCGCGCTCTGCA 184
    ::::: ||||| ::::: ||||| ::::: |||
94  gAlaThrHisValArgAspGlyGlnThrValIleGlyGlnIleLeu 111
    ||||| ||||| ||||| ||||| |||||
185  GGCCATCTCTGTGCCACCGCGCATGCGGTGCCGAGGCGCATGCTGA 234
    ||||| ||||| ||||| ||||| |||||
235  TCGAATCTCGAG 245
    ||||| ||||| ||||| ||||| |||||
111  lGluLeuAsp 114
    ||||| ||||| ||||| ||||| |||||

```

seq_name: gb_gss:AG163349

seq_documentation_block:

LOCUS AG163349 682 bp DNA linear GSS 09-JAN-2002
DEFINITION Pan troglodytes DNA, clone: RP43-030A23.T7, genomic survey
sequence.

ACCESSION

AG163349.1 GI:16693027

VERSION

GSS: (genome survey sequence).

KEYWORDS

Pan troglodytes male lymphocytes DNA, clone:lib:RP43-43 Chimpanzee

SOURCE

Male BAC library clone:RP43-030A23.T7.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

alignment_scores:
Quality: 85.00 Length: 140
Ratio: 1.164 Gaps: 6
Percent Similarity: 52.143 Percent Identity: 30.714

alignment_block:

US-09-484-577a-4 x AG163349 ..

Align seg 1/1 to: AG163349 from: 1 to: 682

```

3 SerLeuSerProArgMetLysSerAlaArgGluValAla..... 16
||||| ||||| ||||| ||||| |||||
83 AGTCTCTACCCAGCGCTGCAATGATTCATGCTCAGCTCTCTAGAG 132
||||| ||||| ||||| ||||| |||||
17 .....ValGlyGlyThrArgAspGlnLeuAlaPheLeuProA 30
||||| ||||| ||||| ||||| |||||
133 CCTCCACCTCTGCGCTTCAATGATTCATGCTCAGCTCTCTAGAG 182
||||| ||||| ||||| ||||| |||||
30 lAlaLeuGluLeuValGluThrProProSerProThrAlaArgLeuThr 46
||||| ||||| ||||| ||||| |||||
183 CTGGGAT.....ACAGATGCTTGCACCATGCGCGCTAATT 220
||||| ||||| ||||| ||||| |||||
47 lAlaAlaLeuLeuAlaAlaLeuPheTYrCysAlaValAlaTrpAlaGlyLe 63
||||| ||||| ||||| ||||| |||||
221 TCTATATTTTGTAGTAGAGCGGCTTTGCCATGTGTTGCCAGCTGCT 270
||||| ||||| ||||| ||||| |||||
63 uGlyArgIle.....AspIleValAlaSerAlaSerArgIleVal 77
||||| ||||| ||||| ||||| |||||
271 TCAACTCTGACCTCATGATCCGCTGCTGCTGCTGCTGCTGCTGCT 320
||||| ||||| ||||| ||||| |||||
77 alProGlyAspArgVal.....LysLeuVal 85
||||| ||||| ||||| ||||| |||||
321 CACTTATAGATAGACTCCAAAGTATATTAAGATTGAGAAACCTGCT 370
||||| ||||| ||||| ||||| |||||
86 GlnProLeu...GluValGlyValValArgAlaThrHisValArgAspGln 101
||||| ||||| ||||| ||||| |||||
371 CAGAACATCTTCAGATACATGATGATTAAGACAGAAATGCACTC 420
||||| ||||| ||||| ||||| |||||
101 yGlnThrValIleValAlaGlyGlnLeu.....LeuIleLeuLeuA 114
||||| ||||| ||||| ||||| |||||
421 CCAAACTCCAGAAAGTGTGAGCTGTGCTGCTGCTGCTGCTGCTGCT 470
||||| ||||| ||||| ||||| |||||
114 sProPheAlaGlyGlyVal 120
||||| ||||| ||||| ||||| |||||
471 GACCTTGGGAGGTAACCTTA 490
||||| ||||| ||||| ||||| |||||

```

seq_name: gb_est:BG274115

seq_documentation_block:

LOCUS BG274115 432 bp mRNA linear EST 21-FEB-2001
DEFINITION WHE2231.G01_N01ZS Aegilops speltoides anther cDNA library Aegilops
speltoides cDNA clone WHE2231.G01_N01, mRNA sequence.

ACCESSION

BG274115.1 GI:13066099

VERSION

KEYWORDS

SOURCE

ORGANISM

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

Submitted (02-AUG-2001) Aseo Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suenho-chou,Tsukumi-Ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimbeg@gsr.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library RP43-43 This BAC
end was generated during the R&D process and may have higher chance
of clone tracking errors.
PRIMERS
Sequencing: T7
LIBRARY
Vector : pBAC3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI.
Location/Qualifiers
1. 682
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="RP43-030A23.T7"
/sex="male"
/cell_type="Lymphocytes"
/clone_lib="RP43-43 Chimpanzee Male BAC library"
183 a 138 c 138 g 221 t 2 others

Fax: 5105595818
Email: oandersn@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20
Seq primer: Stratagene SK primer.
Location/Qualifiers

1. 432
/organism="Aegilops speltoides"
/cultivar="F2 from 2-12-4-8-1-1-1 (1) x PI36909-12-811-(1)"
/db_xref="taxon:4573"
/clone="WHE2231.G01.N01"
/clone_lib="Aegilops speltoides anther cDNA library"
/tissue_type="anther"
/dev_stage="Premeiotic anthers"
/lab_host="E. coli SOLR"

/note="Vector: Lambda uni-ZAP XR, excised phagemid; Site-1: EcoRI; Site-2: XhoI; Plants were grown in a growth chamber at the University of California, Davis (Akhunov). Premeiotic anthers were harvested, total RNA and poly(A) RNA were prepared, from each tissue and then pooled, a cDNA library was made, and the cDNA clones were in vivo excised to give plasmid phagemids in the TU Close Lab (Akhunov, Chin, Choi, Close, Fenton, Kianian, Otto, Simons, Zhang) at the University of California, Riverside. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson Lab (all other authors)."
BASE COUNT 81 a 128 c 134 g 89 t
ORIGIN

alignment_scores:
Quality: 84.00 Length: 67
Ratio: 1.867 Gaps: 3
Percent Similarity: 67.164 Percent Identity: 38.806

alignment_block:
US-09-484-577A-4 x BG274115 ..

Align seg 1/1 to: BG274115 from: 1 to: 432

```

3 SerLeuSerProArgMetLysSerAlaArgIuValAlaAlaAlaGlyI 19
|||||
33 AGCTTGTCCCTTATGGCGCGCAAGACCGTGTGACCAAGCTCG 82
|||||
19 YLSTThArgAspLeuAlaPheLeuProAlaAlaLeuGluLeuAla 36
|||||
83 AAGAACACGCGCTGAGCGACATTCCTCAGCGCGGCTG.....GCCA 126
|||||
36 IuThrProSerProThrAlaArgLeuThrAlaAlaLeuAlaAla 52
|||||
127 GAGGCCGCGCACAGCCAC...AGCGTGAATCCACATTAATTTGGCA 173
|||||
53 Leu.....PheTYCYsAlaValAlaTrrPalaGlyLeuGlyArgI 66
|||||
174 ATGCCCATGCAATCATGTTCAATGCGATGCGGCGCTGCGCGGT 223
|||||
66 e 66
224 G 224

```

seq_name: gb_gss:AQ991716

seq_documentation_block:
LOCUS AQ991716 756 bp DNA linear GSS 14-AUG-2000
DEFINITION Photorhabdus luminescens strain W14 M13 library
ACCESSION AQ991716
VERSION AQ991716.1 GI:9650212
KEYWORDS GSS.
SOURCE Photorhabdus luminescens.
ORGANISM Photorhabdus luminescens
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

REFERENCE
AUTHORS 1 (bases 1 to 756)
TITLE french-Constant R.H., Waterfield,N., Burland,V., Perna,N.T., Daborn,P.J., Bowen,D. and Blattner,F.R.
JOURNAL A genomic sample sequence of the entomopathogenic bacterium
MEDLINE Photorhabdus luminescens W14: potential implications for virulence
COMMENT Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)
20378633
Contact: french-Constant RH
Department of Biology and Biochemistry
University of Bath
South Building, Bath BA2 7AY, UK
Tel: (44) 1225 826621
Fax: (44) 1225 826779
Email: bsr@icbath.ac.uk

FEATURES
source
1. 756
/organism="Photorhabdus luminescens"
/strain="W14"
/db_xref="taxon:29488"
/clone="PLG00179F"
/clone_lib="Photorhabdus luminescens strain W14 M13 library"
/dev_stage="primary phase variant"
/note="Genomic DNA from strain W14 was size selected (1-2 kb) and then cloned into M13 Janus."
BASE COUNT 185 a 205 c 132 g 214 t 20 others
ORIGIN

alignment_scores:
Quality: 84.00 Length: 55
Ratio: 2.270 Gaps: 2
Percent Similarity: 67.273 Percent Identity: 38.182

alignment_block:
US-09-484-577A-4 x AQ991716/rev ..

Align seg 1/1 to reverse of: AQ991716 from: 1 to: 756

```

60 TrrPalaGlyLeuGlyArgIleAspIle...ValAlaSerAlaSerArgI 75
|||||
509 TGGCGACGATTAGCGCGCTGTGAATTAAGCGGTGCGCTTCGCGGTGCT 460
|||||
75 sIleValProGlyAspArgValLysLeuValGlnProLeuGluValGly 92
|||||
459 GGTGTGATGTGTAATCGT...AAACAGTTCAGCCCGCGCAACGGGA 413
|||||
92 alValArgAlaThrHisValArgAspGlyGlnThrValLysAlaGlyIu 108
|||||
412 TTATCAGCCAGATTAAGTTGTTGAGGAGAAACATTAACCGCGGTGAG 363
|||||
109 IleLeuIleGluLeu 113
|||||
362 ACACCTGTCACAGCTC 348

```

seq_name: gb_est2:BG758850

seq_documentation_block:
LOCUS BG758850 973 bp mRNA linear EST 15-MAY-2001
DEFINITION 602713239F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4653577 5', mRNA sequence.
ACCESSION BG758850
VERSION BG758850.1 GI:14069503
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
1 (bases 1 to 973)
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
NIH-MGC <http://mhc.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: InCyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/MLML at: <http://image.llnl.gov>
Plate: LHCML699 row: n column: 10
High quality sequence stop: 2.

FEATURES

Location/Qualifiers
1..973

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4853577"
/clone_lib="NIH-MGC.48"
/tissue_type="Primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/note="Organ: B-cells; Vector: pORF7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGGAG(G). Size-selected for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC library."

BASE COUNT 232 a 261 c 279 g 199 t 2 others
ORIGIN

alignment_scores:

Quality: 83.50 Length: 77
Ratio: 1.815 Gaps: 3
Percent Similarity: 59.740 Percent Identity: 35.065

alignment_block:

US-09-484-577A-4 x BG758850 ..

Align seg 1/1 to: BG758850 from: 1 to: 973

```

49 LeuLeuAlaAlaLeuPheTyrCysAlaValAlaTrpAlaGlyLeuGlyAar 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
473 CTGCTAGGGAGGAGAACTTTACTCTTAGAGAGGAGTGTGCTCTGCGAG 522
65 gtleaspileval.....AlaserAlaserArgLysilevalProGlyA 80
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
523 GTTGTCACACTGTTATCCACGCGAGCGTATGCGTGTGGGGTGGGT 572
80 spArg.ValLysLeuValGlnProLeuGluValGlyValAlaTrpAlaTh 96
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
573 CCGGTGGCATTGTTGTCTCCGCAATTCGCTGGGCTCTTACGCGTTCG 622
96 rhtsValArgAspGlyGlnThrValLysAlaGlyLuleuLeuIleGluL 113
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
623 GAGAGTCCGGGACGGGGGCGATCCCGCT..... 650
113 euAspProPheAlaGlyLysValAlaSpVal 122
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
651 .....GAGTATTCGCGCGAGTGTAGTGC 674
seq_name: gb_est2:BJ189099

```

seq_documentation_block: 456 bp mRNA linear EST 24-JAN-2002
LOCID BJ189099
DEFINITION BJ189099 normalized full length cDNA library, chloronemata,

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

TITLE
JOURNAL
COMMENT

caulonemata and malformed buds *Physcomitrella patens* subsp. *patens*
cDNA clone pphb45c13 5', mRNA sequence.
BJ189099.1 GI:18357040
EST.
Physcomitrella patens subsp. *patens*.
Physcomitrella patens subsp. *patens*.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
Bryopsida; Funariidae; Funariales; Funariaceae; *Physcomitrella*.
1 (bases 1 to 456)
Fujita,T., Shin-I,T., Seki,M., Kamiya,A., Uchiyama,I., Nishiyama,T.,
Carninci,P., Hayashizaki,Y., Shinozaki,K., Kohara,Y. and Hasebe
M.
Comparison of the moss *Physcomitrella patens* genome with flowering
plants genome
Unpublished (2002)
Contact: Tadasi Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp
A backbone of the vector is pBluescript II, that was in vivo
excised from a modified lps phage vector (Mo bi Tec, Germany). XhoI
digested-5' end of cDNA is ligated to SalI site of the vector, and
the BamHI digested-3' end including poly-A tail is ligated to BamHI
site of the vector. cDNA insert could be amplified with
conventional T7 and T3 primers. This normalized full-length cDNA
library was generated basically according to the method described
in Genome Research 10, 1617-1630 (2000). Carninci, P. et al.
Protonemata were blended by the POLYTRON, and then cultivated on
the BCD medium containing 0.5mM BA (benzylaminopurine) for 8 to 13
days under the continuous light.
Location/Qualifiers
1..456
/organism="Physcomitrella patens subsp. patens"
/db_xref="taxon:145481"
/clone="pphb45c13"
/clone_lib="normalized full length cDNA library,
chloronemata, caulonemata and malformed buds"
/tissue_type="mixture of chloronemata, caulonemata and
malformed buds"

BASE COUNT 97 a 97 c 129 g 133 t
ORIGIN

alignment_scores:

Quality: 82.00 Length: 93
Ratio: 1.608 Gaps: 3
Percent Similarity: 54.839 Percent Identity: 26.882

alignment_block:

US-09-484-577A-4 x BJ189099 ..

Align seg 1/1 to: BJ189099 from: 1 to: 456

```

38 ProProseProthrmAlaArgLeu..... 45
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
25 CCCCCCCTCCGACCTCCGCGCTGTACTGTCTGCGAGAGAC 74
46 .ThraAlaLeuLeuAlaLeuPheTyrCysAlaValAlaTrpAlaG 62
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
75 TACAGTTTGTACATCCCTGACATTGTGCTGCGACGTT.....C 115
62 LysLeuGlyArgLysPheAlaAlaSerAlaSerArgLysilevalPro 78
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
116 ATTTAGTCGAGATTGACTTGCTGTCGAGCGCTGAGACCTTGTAGAC 165
79 GlyAspArgValLysLeuValGlnProLeuGluValGlyVal..... 92
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
166 GTGAGAGATACACTTTCGCCACCCGTCGAGAACTTGGACGCGGTGTGA 215

```



```

67 .....AspIleValAlaSerAlaLeu 73
263 GTGCTTTTCATGCGACCAAGCATGCTCCGACATGAGCATCTGTGGC 312
73 rArgLysIleValProGly.....AspArgValLysLeuValG 86
313 GCGCTCTGTGTGCCCCCGCATTTTCGAGAGGATCTCTTCAGCTTTNAC 362
86 InProLeuGluValGlyValAlaValAlaThrHisValAlaArgSpolGly 102
363 ACACGCTCCGATGTGCTGTGCATCAGCCCTCAGACCAAGCTCAGAGGAG 412
103 ThrValLysAlaGlyLysIleLeuIleGluLeuAspProphealGly 119
413 TCCAGCTTCGAGGCGCCAGCTNMAAGCCTGTGCATNTACAGGCTGTC 462
119 yValAsp 121
463 TGTGAC 469
seq_name: gb_est1:AL514080

```

```

seq_documentation_block:
LOCUS AL514080 907 bp mRNA linear EST 13-FEB-2001
DEFINITION AL514080 LFI_NFL006_PL2 Homo sapiens cDNA clone CIOBA004ZE11 5
ACCESSION AL514080
VERSION AL514080.1 GI:12777574
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 907)
Li W.B., Gruber,C., Jesse,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
COMMENT
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segref@genoscope.cns.fr, web : www.genoscope.cns.fr.
FEATURES
Source
1..907
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CIOBA004ZE11"
/clone_lib="LFI_NFL006_PL2"
/tissue-type="placenta"
/note="Vector: PCWSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the PCWSPORT 6
vector. Library was normalized. Library was constructed by
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : liliang@lifestech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 192 a 236 c 273 g 194 t 12 others
ORIGIN

```

```

alignment_scores:
Quality: 80.00 Length: 69
Ratio: 1.860 Gaps: 1
Percent Similarity: 62.319 Percent Identity: 36.232

```

alignment_block:

us-09-484-577a-4 x AL514080/rev ..

Align seq 1/1 to reverse of: AL514080 from: 1 to: 907

```

40 SerProThrAlaArgLeuThrAlaLeuLeuAlaAlaLeuPheTyr 56
||||| :||| ||||| :||| :|||
280 TCGCGGAGACCTGCTGCTGCTGCGAGCTGCTGCTGCTGCTGCTGCTG 231
56 sAlaValAlaTrpAlaLeuGlyArgGlyLeuAspIleValAlaSerAla 73
||||| :||| :||| :||| :||| :|||
230 TGTCTGGGAMTWTGTTACTCTGCGAGATTGCCCTGCTGCTGAGAGTA 181
73 eArgLysIleValProGlyAspArgValLysLeuVal.....GlnPro 87
180 GTGTAAGAGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 131
130 CTGCTGCTGCGAATGATGCGAGTCTTCACACGACCTCCGTGAGCCG 81
104 lLysAla 106
80 CCGGCA 74

```

```

seq_name: gb_est2:BE481913
seq_documentation_block:
LOCUS BE481913 360 bp mRNA linear EST 28-AUG-2000
DEFINITION 167524 BARC 5BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BE481913
VERSION BE481913.1 GI:9601446
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE
1 (bases 1 to 360)
Sonstegard,T.S., Capuco,A.V., Van Tassel,C.P., Ashwell,M.S. and
Wells,K.D.
Mapping of Expressed Sequence Tags from a normalized bovine mammary
gland cDNA library
Unpublished (2000)
JOURNAL
COMMENT
Contact: Sonstegard TS
USDA, ARS, Beltsville Agricultural Research Center
Bldg. 200 Rm 2A, Beltsville, MD 20705, USA
Tel: 301 504 8416
Fax: 301 504 8414
Email: tads@psi.barc.usda.gov
Single pass sequencing. Bases called and alt-trimmed with phred
v0.980904.e. Vector identified by cross-match with the -minscore 18
PCR Primers
and -mismatch 12 options.
FORWARD: AGGAACAGCTATGACAT
BACKWARD: GTTTCACAGTACAGCAGC
Plate: 19 row: 0 column: 4
Seq primer: ATTGAGTGACACTATG.
FEATURES
Source
1..360
location/Qualifiers
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="BARC 5BOV"
/tissue-type="pooled"
/lab_host="DH10B"
/note="Vector: PCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled mRNA isolated from mammary
tissues at eight physiological, developmental, and disease
states.."
BASE COUNT 73 a 123 c 107 g 57 t
ORIGIN

```

```

alignment_scores:
Quality: 79.50 Length: 84
Ratio: 1.691 Gaps: 3
Percent Similarity: 55.952 Percent Identity: 32.143

```

alignment_block:

US-09-484-577A-4 x BE481913/rev ..

Align seg 1/1 to reverse of: BE481913 from: 1 to: 360

```

39 ProserProthrAlaArgLeuThrAlaLeuAlaLeuAlaLeuPheTy 55
   |||||.....: |||.....: |||.....:
319 CCGAGTCGGGTTCAGGTTCCGCCGAAATCTGTGCCGTCTCTGT 270
   :|||.....: |||.....: |||.....:
55 rcysAlaValAlaTrpAlaGlyLeuGlyArgLeuAlaLeuAlaSer 72
   :|||.....: |||.....: |||.....:
269 AGATTCTGTGTGTGGATCCCACTCCGCCCTCTCTCCATCCAC 220
   :|||.....: |||.....: |||.....:
72 LaserArgLysIleValProGlyAspArgValLysLeuValGlnPro 88
   :|||.....: |||.....: |||.....:
219 GGCACCCCGCGGTTCCTGGATTCCGGCGCTCCGTGCAACCGTGT 170
   :|||.....: |||.....: |||.....:
89 GluValGlyValAlaArgAlaThrHisValArgAspGlyGlnThrVal 105
   |||.....: |||.....: |||.....:
169 GAACCTGGTGTCTGTC.....CACGTAGCCGACGCGCAT.....GA 135
   :|||.....: |||.....: |||.....:
105 sAlaGlyGluIleLeuIleLeuAspProPheAlaGlyValAspV 122
   :|||.....: |||.....: |||.....:
134 AGCGGGG.....CTCCCGAGGGCGCGCGACACGCGGTGTAGAAA 91
   :|||.....: |||.....: |||.....:
122 al 122
   :|||.....: |||.....: |||.....:
90 TA 89

```

seq_name: gb_est2:D27997

seq_documentation_block:

LOCUS D27997 394 bp mRNA linear EST 20-NOV-1995
 DEFINITION CELR007EIF yuji kohara unpublished cDNA Caenorhabditis elegans cDNA
 clone yk7el 5', mRNA sequence.

ACCESSION D27997

VERSION D27997.1 GI:522753

KEYWORDS EST

SOURCE

ORGANISM

Caenorhabditis elegans.

Caenorhabditis elegans.

Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea

1 (bases 1 to 394)

Kohara, Y., Mitsuki, H., Nishigaki, A., Motohashi, T., Sugimoto, A. and

Tabara, H.

Toward an expression map of the C. elegans genome

Unpublished (1994)

Contact: Yuji Kohara

Genome Biology Lab.

National Institute of Genetics

Yata 1111, Mishima, Shizuoka 411, Japan

Tel: 81-559-81-6854

Fax: 81-559-81-6855

Email: ykohara@lab.nig.ac.jp.

Location/Qualifiers

1..394

/organism="Caenorhabditis elegans"

/strain="CB1489 him-8(e1489)"

/db_xref="taxon:6239"

/clone="yk7el"

/clone_lib="Yuji Kohara unpublished cDNA"

/sex="hermaphrodite, male"

/tissue_type="whole animal"

/dev_stage="varied"

BASE COUNT 112 a 93 c 85 g 102 t 2 others

ORIGIN

alignment_scores:

Quality: 79.50 Length: 56
 Ratio: 2.650 Gaps: 1
 Percent Similarity: 53.571 Percent Identity: 33.929

alignment_block:

US-09-484-577A-4 x D27997 ..

Align seg 1/1 to: D27997 from: 1 to: 394

```

75 LysIleValProGlyAspArgValLysLeuValGlnProLeuGluVal.. 90
   |||||.....: |||.....: |||.....:
86 AAATCAAGAGGAGCGACAGAGTTACCAAGACACCCCTATTGTGTCT 135
   :|||.....: |||.....: |||.....:
91 .....GlyValV 93
   :|||.....: |||.....: |||.....:
136 TTCCGCCCATGAAATGGAGATGTCATGACCTCTCAATTGCTGGAACG 185
   :|||.....: |||.....: |||.....:
93 aLArgAlaThrHisValArgAspGlyGlnThrValLysAlaGlyGluIle 109
   :|||.....: |||.....: |||.....:
186 TCAAGCTATTCATGACACACACAGAACCAATGTATGCGGAGACTTG 235
   :|||.....: |||.....: |||.....:
110 LeuIleGluLeuAspPro 115
   :|||.....: |||.....: |||.....:
236 GTAGTTGAAGTTGACCA 253

```

Tue Jul 16 05:46:34 2002

us-09-484-577a-4.rst

Page 10

Percent Similarity: 71.845 Percent Identity: 49.515

Alignment block:

US-09-484-577A-4 x AA254335

Align seg 1/1 to: AA254335 from: 1 to: 1428

```

20 LysThrArgAspGluLeuAlaPheLeuProAlaAlaLeuGluIleValG1 36
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
94 CGCAGCGCGGAGAACAGCGCTTTGGCCCGCATTTGGAACTGACCGCA 143
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
36 uThrProSerProThrAlaArgLeuThrAlaAlaLeuAlaAlaL 53
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
144 TACGCGGCTCTGCGCTCGCAATGGCGCGCGCTTTATATATGCGGT 193
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
53 eupheryCysAlaValAlaTrpAlaGlyLeuGlyArgGlyIleVal 69
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
194 TTGCGCTTTGGCTTTGTGTGTGTCGTCGCGCAAAATCGATATGTG 243
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
70 AlaSerAlaSerArgGlyIleValProGlyAspArgValIysLeuValG1 86
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
244 GCGCGGCTTGGCGGCAAAACGCTCGCGCGCGCGCAAAACCATCCA 293
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
86 nProLeuGluValGlyValAlaArgAlaThrHisValArgAspGlyLnt 103
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
294 GCGCGTGAACAGCGCGGTGTTAAGCGGTACATGTCGCGCGCGGACG 343
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
103 hrValIysAlaGlyIleLeuIleGluLeuAspProPheAlaGlyGly 119
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
344 ATGTGAACAGCGGAGAAACGCTGCGGCACTGAGCGCTGTGGGAACAGAC 393
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
120 ValAspVal 122
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
394 AGCGATGTG 402
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::

```

seq_name: /SID1/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:AA81532

seq_documentation_block:

ID AA81532 standard; DNA; 24158 BP.

AA81532;

04-DEC-2000 (first entry)

N. meningitidis partial DNA sequence gnm_79 SEQ ID NO:79.

Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
antigen; vaccine; diagnosis; infection; antibacterial; identification;
Meningococcus B; Memb; ds.

Neisseria meningitidis.

MO200022430-A2.

20-APR-2000.

08-OCT-1999; 99WO-US23573.

09-OCT-1998; 98US-0103794.

30-APR-1999; 99US-0132068.

(CHIR) CHIRON CORP.

Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
Maignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
Rappuoli R, Pizza M;

WPI: 2000-318079/27.

Isolated nucleotide sequences of Neisseria meningitidis which can be
used in the diagnosis and treatment of N. meningitidis infection and
other Neisserial infections, for example, N. gonorrhoea -

PS Claim 7: Page 1464-1471; 1760pp; English.

XX The present invention describes methods of obtaining immunogenic
CC proteins from Neisseria genomic sequences. AA81453 to AA82414
CC represent specifically claimed Neisseria meningitidis genomic DNA
CC sequences; AA81260 to AA81303 and AA825620 to AA825663 represent
CC Neisseria DNA sequences and their corresponding proteins; AA81254 to
CC AA81259 and AA81304 to AA81321 represent PCR primers used in the
CC isolation of Neisseria meningitidis DNA sequences; and AA81322 to
CC AA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC sequences, which are all used in the exemplification of the present
CC invention. The nucleic acid sequences, protein sequences, and antibodies
CC against them, can be used in the manufacture of a composition. The
CC composition can be used as a medicament (or in the manufacture of a
CC medicament) for treating, preventing or diagnosing infection due to
CC Neisserial bacteria. For example, some of the identified proteins could
CC be components of vaccines against Meningococcus B; against all serotypes;
CC and/or against all pathogenic Neisseriae. Identification of biological probes,
CC particularly organism-specific probes. Attempts to make efficacious
CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC Multivalent vaccines have also been tried but none have successfully
CC overcome antigenic variability. The provision of further, complete
CC sequences may provide an opportunity to identify secreted or surface
CC exposed proteins that may be presumed targets for the immune system and
CC which are not antigenically variable or at least more conserved than
CC other more variable regions.

Sequence 24158 BP; 5559 A; 6297 C; 6241 G; 6061 T; 0 other;

alignment_scores:

Quality: 238.00 Length: 103
Ratio: 3.216 Gaps: 0
Percent Similarity: 71.845 Percent Identity: 49.515

alignment_block:

US-09-484-577A-4 x AA81532

Align seg 1/1 to: AA81532 from: 1 to: 24158

```

20 LysThrArgAspGluLeuAlaPheLeuProAlaAlaLeuGluIleValG1 36
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
17976 CGCAGCGCGGAGAACAGCGCTTTGGCCCGCATTTGGAACTGACCGCA 18025
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
36 uThrProSerProThrAlaArgLeuThrAlaAlaLeuAlaAlaL 53
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
18026 TACGCGGCTCTGCGCTCGCAATGGCGCGCGCTTTATATATGCGGT 18075
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
53 eupheryCysAlaValAlaTrpAlaGlyLeuGlyArgIleAspIleVal 69
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
18076 TTGCGCTTTGGCTTTGTGTGTGTCGTCGCGCAAAATCGATATGTG 18125
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
70 AlaSerAlaSerArgGlyIleValProGlyAspArgValIysLeuValG1 86
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
18126 GCGCGGCTTGGCGGCAAAACGCTCGCGCGCGCGCAAAACCATCCA 18175
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
86 nProLeuGluValGlyValAlaArgAlaThrHisValArgAspGlyLnt 103
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
18176 GCGCGTGAACAGCGCGGTGTTAAGCGGTACATGTCGCGCGGACG 18225
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
103 hrValIysAlaGlyIleLeuIleGluLeuAspProPheAlaGlyGly 119
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
18226 ATGTGAACAGCGGAGAAACGCTGCGGCACTGAGCGCTGTGGGAACAGAC 18275
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
120 ValAspVal 122
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
18276 AGCGATGTG 18284
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::

```

seq_name: /SID1/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:AAF21611

seq_documentation_block:

ID AAF21611 standard; DNA; 34980 BP.


```

alignment_scores:      238.00      Length:      103
                      Quality:      103
                      Ratio:      3.216      Gaps:      0
Percent Similarity:    71.845      Percent Identity: 49.515

alignment_block:
US-09-4884-577A-4 x AAF21611/rev ..

Align seg 1/1 to reverse of: AAF21611 from: 1 to: 349860

```

[illegible]

CC sequences which overlap each other at the beginning and end of each
 CC sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
 CC the beginning of AAF21607, the last 49980 bp of AAF21607 are repeated at
 CC the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
 CC Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
 CC AAF21606 represent PCR primers which are used in the exemplification of
 CC the present invention. The NMB genome and fragments from it have
 CC antibacterial activity, and can be used in vaccines and gene therapy.
 CC Neisseria nucleic acids, proteins and/or antibodies which binds to the
 CC proteins can be used in compositions for treating or preventing infection
 CC due to Neisserial bacteria or as a diagnostic reagent for detecting the
 CC presence of Neisserial bacteria or of antibodies raised to Neisserial
 CC bacteria. Computers, computer memory, computer storage medium or computer
 CC databases can be used in a search to identify open reading frames (ORFs)
 CC or coding sequences within the NMB genome. The DNA sequences provide
 CC further opportunities to find antigenic or immunogenic proteins which are
 CC more effective in vaccines than the outer membrane proteins currently
 CC used.

Sequence 349980 BP; 86473 A; 95646 C; 85908 G; 81953 T; 0 other;

alignment_scores:

Quality: 238.00 Length: 103
 Ratio: 3.216 Gaps: 0
 Percent Similarity: 71.845 Percent Identity: 49.515

alignment_block:

US-09-484-577A-4 x AAF21612/rev ..

Align seg 1/1 to reverse of: AAF21612 from: 1 to: 349980

```

20 LysThrArgAspGluLeuAlaPheLeuProAlaAlaLeuGluIleValG1 36
   ::::: ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
23829 CGACGCGCGAGAACAGCGGCTTTTGGCCGCGCATTTGGAACTGACCGA 23780
   ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
36 uThrProSerProThrAlaArgLeuThrAlaAlaLeuAlaLeuAlaL 53
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
23779 TACGCGCGCTCTGCTCCGCAATGGGCGCGCTTTTATTATATGCGCT 23730
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
53 euPheTyrCysAlaValAlaTrpAlaGlyLeuGlyArgIleAspIleVal 69
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
23729 TTGCGCCTTTTGGCTTTGTTGGTCTGCTGCGCAAAATGATATGTG 23680
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
70 AlaSerAlaSerArgIleValProGlyAspArgValLysLeuValG1 86
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
23679 GCGCGCGCTTGGCGCAAAAGGTGCGCGCGCGCGCAAAACCAATCCA 23630
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
86 nProLeuGluValGlyValAlaArgAlaThrHisValArgAspGlyG1 103
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
23629 GCCGCTGGAACGGCGGTGTTAAGCGGTACATGTGCGCGACGCGCAGC 23580
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
103 hrValLysAlaGlyIleLeuIleLeuLeuLeuAspProPheAlaGly 119
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
23579 ATGTGAACAGGAGAAACGCTGCGGAACTGAGAGCTGTGGAAACAGAC 23530
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
120 ValAspVal 122
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
23529 AGCGATGTG 23521

```

seq_name: /SIDSI/gcgdata/hold-geneseg/geneseg-emb1/NA2000.DAT:AAA81489
 seq_documentation_block:
 ID AAA81489 standard: DNA; 837096 BP.

AC AAA81489;
 XX
 XX 04-DEC-2000 (first entry)
 XX
 XX N. meningitidis partial DNA sequence gmm_37 SEQ ID NO:37.
 DE
 XX Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
 KW antigen; vaccine; diagnosis; infection; antibacterial; identification;

KW Meningococcus B; MenB; ds.
 XX
 XX Neisseria meningitidis.
 OS
 XX
 XX WO200022430-A2.
 PN

XX 20-APR-2000.
 XX
 XX 08-OCT-1999; 99WO-US23573.
 XX
 XX 09-OCT-1998; 98US-0103794.
 PR
 XX 30-APR-1999; 99US-0132068.
 XX

PA (CHIR) CHIRON CORP.

PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
 PI Masigiani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
 PI Rappuoli R, Pizzi M;
 XX
 XX WPI: 2000-318079/27.

XX Isolated nucleotide sequences of Neisseria meningitidis which can be
 XX used in the diagnosis and treatment of N. meningitidis infection and
 XX other Neisserial infections, for example, N.gonorrhoea -
 PS Claim 7: Page 629-865; 1760pp; English.

CC The present invention describes methods of obtaining immunogenic
 CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414
 CC represent specifically claimed Neisseria meningitidis genomic DNA
 CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
 CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to
 CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
 CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
 CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
 CC sequences, which are all used in the exemplification of the present
 CC invention. The nucleic acid sequences, protein sequences, and antibodies
 CC against them, can be used in the manufacture of a composition. The
 CC composition can be used as a medicament (or in the manufacture of a
 CC Neisserial bacteria). For example, some of the identified proteins due to
 CC be components of vaccines against Meningococcus B; against all serotypes;
 CC and/or against all pathogenic Neisseriae. Identification of sequences
 CC from the bacterium will also facilitate production of biological probes,
 CC particularly organism-specific probes. Attempts to make efficacious
 CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
 CC Multivalent vaccines have also been tried but none have successfully
 CC overcome antigenic variability. The provision of further, complete
 CC sequences may provide an opportunity to identify secreted or surface
 CC exposed proteins that may be presumed targets for the immune system and
 CC which are not antigenically variable or at least more conserved than
 CC other more variable regions.
 XX
 XX Sequence 837096 BP; 207534 A; 227065 C; 205215 G; 197280 T; 2 other;

alignment_scores:

Quality: 238.00 Length: 103
 Ratio: 3.216 Gaps: 0
 Percent Similarity: 71.845 Percent Identity: 49.515

alignment_block:

US-09-484-577A-4 x AAA81489/rev ..

Align seg 1/1 to reverse of: AAA81489 from: 1 to: 837096

```

20 LysThrArgAspGluLeuAlaPheLeuProAlaAlaLeuGluIleValG1 36
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
388600 CGACGCGCGAGAACAGCGGCTTTTGGCCGCGCATTTGGAACTGACCGA 388551
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
36 uThrProSerProThrAlaArgLeuThrAlaAlaLeuAlaLeuAlaL 53
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
388550 TACGCCGCTCTGCTCCGCAATGGGCGCGCTTTTATTATATGCGCT 388501

```

	CC	the invention can be used as vaccines, as diagnostic reagents, and as
	CC	immunogenic compositions. The polypeptides can be used in the
	CC	manufacture of medicaments for treating or preventing infection due to
	CC	Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
	CC	presence of Neisseria bacteria, or to raise antibodies. They may also
	CC	be used to screen for agonists or antagonists, which may themselves
	CC	have use as antibacterial agents. The polynucleotides of the invention
	CC	may also be used in gene therapy protocols.
	XX	
	XX	Sequence 1305 BP; 303 A; 301 C; 431 G; 270 T; 0 other:
		alignment_scores:
		Quality: 235.00 Length: 103
		Ratio: 3.219 Gaps: 0
		Percent Similarity: 70.874 Percent Identity: 49.515
		alignment_block:
		US-09-484-577A-4 x AA54336 ..
		Align seg 1/1 to: AA54336 from: 1 to: 1305
		20 LysThrArgAspGluLeuAlaPheLeuProHalaIeuGluIleValGl 36
		:::: :::: :::::
		94 CGCAGCGGGGAGAAGACAGCGCTTTTGGCCGCATTTGGAACTGACCGA 143
		36 uHrProProSerProThraIArgLeuThraAlaIaleuAlaIal 53
		::: :::: :::: :::::
		144 TACGCCCGGTCTGCCGCTCCGAATAAGGCGGCGGCTTTATTATMGCGGT 193
		53 euPheTrcYsaIaValaIatrrPalaglyLeuGlyArgIIeaSplIaVal 69
		::: ::: :::: ::: :::: ::: :::: :::
		194 TTGCGCTTTTGCTTGTGTGGTCCTGCTGGCAAAATGATATTGTG 243
		70 AlaSerAlaSerArgLysIleValProGlyAspArgValLysIleuValGl 86
		244 GCGGCGGCTTCGGCAAAAGSTGTGCGGCGGCGCACCAAAACCATCCA 293
		86 nProIeuGluValIGlyValIaIArgAlaThHisValaIArgAspGIyInt 103
		::: :::: ::: ::: ::: ::: ::: ::: ::: :::
		294 GCGCGTCAAGACGTGCTGTTAAAGCGGTATCATGTGCGCGGCGAGC 343
		103 hrValLysLaGIgIuIleuIIeGIleuAspPrPheAlaGlyely 119
		::: :::
		344 ATGTGAACAAGGAGAACCTGCGGCAACTGAGGCTGTGGGAACAGAC 393
		120 ValAspVal 122
		394 AGCGATGTG 402
		seq_name: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NM1996.DAT:AA137329
		seq_documentation_block:
		ID AA137329 standard; DNA: 28804 BP.
		AC AA137329;
		XX
		DT 30-NOV-1996 (first entry)
		XX
		Sphingian biosynthetic gene region.
		DE
		KW Sphingari; polysaccharide; spsB gene; glucosyl-IP-transferase; ds.
		XX
		OS Sphnomonas strain S88 (ATCC 31554).
		XX
		key Location/Qualifiers
		FH complement (1942..1944)
		FT /*tag= a
		FT /codon_start= 1942..1944
		FT /note= "spsB gene putative initiation codon"
		FT complement (3311..3313)
		FT /*tag= b
		CDS /codon_start= 3311..3313
		FT

```

FT      /note= "spss gene putative initiation codon"
FT      /tag= c
FT      /codon_start= 5323..5325
FT      /note= "spss gene putative initiation codon"
FT      /tag= d
FT      /codon_start= 5526..5528
FT      /note= "spss gene putative initiation codon"
FT      /tag= e
FT      /codon_start= 7076..7078
FT      /note= "spss gene putative initiation codon"
FT      /tag= f
FT      /codon_start= 7588..7590
FT      /note= "spss gene putative initiation codon"
FT      /tag= g
FT      /codon_start= 8643..8645
FT      /note= "spss gene putative initiation codon"
FT      /tag= h
FT      /codon_start= 10938..10940
FT      /note= "spss gene putative initiation codon"
FT      /tag= i
FT      /codon_start= 11569..11571
FT      /note= "spss gene putative initiation codon"
FT      /tag= j
FT      /codon_start= 12886..12888
FT      /note= "spss gene putative initiation codon"
FT      /tag= k
FT      /codon_start= 15165..15167
FT      /note= "spss gene putative initiation codon"
FT      /tag= l
FT      /codon_start= 15883..15885
FT      /note= "spss gene putative initiation codon"
FT      /tag= m
FT      /codon_start= 16748..16750
FT      /note= "spss gene putative initiation codon"
FT      /tag= n
FT      /codon_start= 21082..21084
FT      /note= "spss gene putative initiation codon"
FT      /tag= o
FT      /codon_start= 21706..21708
FT      /note= "spss gene putative initiation codon"
FT      /tag= p
FT      /codon_start= 23238..23240
FT      /note= "spss gene putative initiation codon"
FT      /tag= q
FT      /codon_start= 24113..24115
FT      /note= "spss gene putative initiation codon"
FT      /tag= r
FT      /codon_start= 24683..24685
FT      /note= "spss gene putative initiation codon"
FT      /tag= s
FT      /codon_start= 25744..25746
FT      /note= "spss gene putative initiation codon"
FT      /tag= t
FT      /codon_start= 27534..27536
FT      /note= "spss gene putative initiation codon"

```

```

FT      CDS      27747
FT      /tag= u
FT      /codon_start= 27747..27749
FT      /note= "urfs3 gene putative initiation codon"
FT      EP2728841-A2.
FT      28-AUG-1996.
FT      24-JAN-1996; 96EP-0300467.
FT      24-JAN-1995; 95US-0377440.
FT      (SHIN-) SHINETSU BIO INC.
FT      (SHIE) SHINETSU CHEM CO LTD.
FT      Armentrout RW, Mikolajczak M, Pollock TJ, Thorne L;
FT      Yamazaki M;
FT      WPI: 1996-386292/39.
FT      P-PSDB: AAW03997.
FT      New isolated DNA from Sphingomonas sp. - used for transforming
FT      recipient bacteria to obtain hyper-producers of sphing
FT      polysaccharide(s).
FT      Claim 32; Page 56-70; 105pp: English.
FT      A 28.8 kb chromosomal fragment of Sphingomonas strain S88 was
FT      isolated on the basis of its ability to restore sphing
FT      biosynthetic capability to Sphingomonas mutant S88m260. It
FT      contains 23-25 genes, including sps genes coding for biosynthesis of
FT      the polysaccharide sphingon, the genes coding for dTDP-(L)thiamose
FT      biosynthesis, atpB genes coding for a transport function and some
FT      CC was identified that is believed to code for glucosyl IP-transferase
FT      CC (AAW03997), an enzyme catalyzing the first step of assembly of
FT      CC sphingon carboxylates. DNA fragments of 588 can be inserted into
FT      CC a vector in multiple copies and used to produce engineered bacteria
FT      CC that are hyper-producers of sphingon.
FT      SQ      Sequence 28804 BP; 4974 A; 9806 C; 9228 G; 4796 T; 0 other:

```

```

alignment_scores:
  Quality: 215.00      Length: 97
  Ratio: 2.905        Gaps: 0
  Percent Similarity: 76.289      Percent Identity: 45.361

```

alignment_block:

US-09-484-577A-4 x AAT37329/rev ..

Align seg 1/1 to reverse of: AAT37329 from: 1 to: 28804

```

27 PheLeuProAlaAlaLeuGluIleValGluThrProSerProThrAl 43
18808 TTTCTGCGCGCGCGCGCTCGATCATCGACGTCGCGTCCGCCACCCG 18759
43 AArgLeuThrAlaAlaLeuAlaLeuPheThyTyrCysAlaValAla 60
18758 GCGCTCACCGCGCGCGGTGATGCGCGCGCTCGATCACCGCCCT 18709
60 TPAlaGlyLeuGlyArgIleAspIleValAlaSerAlaSerArgIle 76
18708 GCGTGCACATCGCGCGCTCGATGATGCGCGCGCGCGCGCGATC 18659
77 ValProGlyAspArgValGlyLeuValGluProLeuGluValGlyVal 93
18658 GCCCGATCGCGCGCGATGATGATGATGATGATGATGATGATGAT 16609
93 LArgAlaThrHisValArgAspGlyGlnThrValGlyAlaGlyGlu 110
18608 CCGGCGATTTCTGTTGCGCGCGCGCGCGCGCGCGCGCGCGCG 18559

```


Quality: 215.00 Length: 97
 Ratio: 2.905 Gaps: 0
 Percent Similarity: 76.289 Percent Identity: 45.361

Alignment block:

US-09-484-577A-4 x AAV99812/rev ..

Align seg 1/1 to reverse of: AAV99812 from: 1 to: 28804

```

27 PheleuproAlaAlaLeuGluIleValGluThProProserProthral 43
|||||
18808 TTCTGCGCGCGCGCTCGAGATCATCGAACCTCCGTGTGCCACCGC 18759
43 aArgLeuThAlaAlaLeuLeuAlaLeuPheTYrCysAlaValAla 60
|||||
18758 GCGGCTTACCGCGCGGTGATGTGGCGGCGCTGCGATCACCACCGCT 18709
60 rPalagLeuGluArgIleAspIleValAlaSerAlaSerArglySile 76
|||
18708 GCGTGGCAATCGCGCGCGGTGAGAGTGTGGCGGCGGACGAGCGCGCTATC 18659
77 ValProGlyAspArgValLysLeuValGlnProLeuGluValGlyValVa 93
|||||
18658 GCCCGCATCGCGCGACCAAGATCGTCCGCCGGAAGCGGTATCCT 18609
93 lArgAlaThrHisValArgAspGlyGlnThrValLysAlaGlyGluIle 110
|||||
18608 CCGGCGCATTCGTGTGGCGGAGGCGAGAGTTTCAGAGAGGCGCGAGTGC 18559
110 euIlleGluLeuAspProphealagIlyValAspValAla 123
|||||
18558 TGTATCACGCTCGATCCACCATGTGCGGCGGCGGAGCGCGCGC 18518

```

seq_name: /SISL1/gcdata/hold-geneseq/geneseq-emb1/NA1999.DAT.AAV81474

seq_documentation_block:

ID AAV81474 standard; DNA: 28804 BP.

AC AAV81474;

DT 16-MAR-1999 (first entry)

DE Chromosomal fragment of *Sphingomonas* strain 88 genome.

KW Chromosome; sphingane; bacterium; polysaccharide; polymer; additive; food;

KM glycosyl-C55-isoprenylphosphate transferase; textile; cosmetic; paper;

KW paint; cement; viscosity; adhesive; petroleum; chemical; ds.

XX *Sphingomonas* sp.

PH Key Location/Qualifiers

FT CDS 21706..23118 /*tag- a

FT /product= "spsb protein"

PN US854034-A.

PD 29-DEC-1998.

PF 24-JAN-1996; 96US-0592874.

PR 24-JAN-1996; 96US-0592874.

PR 24-JAN-1995; 95US-0377440.

PA (SHIN-) SHIN-ETSU CO INC.

PA (SHIE) SHINETSU CHEM CO LTD.

PI Armentrout RW, Mikolajczak M, Pollock TJ, Thorne L;

XX Yamazaki M,

DR MPI: 1999-094809/08.

DR P-PSDR: AAV67750.

PT Production of sphingane polysaccharide products - by introducing DNA
 PT from sphingane-producing *Sphingomonas* species in multiple copies into
 PT recipient *Sphingomonas* sp.

PS Example 16; Fig 14A-K; 66pp; English.

CC This sequence represents a chromosomal fragment of the *Sphingomonas* sp.
 CC strain 88 genome. DNA from this sphingane-producing bacterium can be
 CC used to increase production of sphingane polysaccharides in other
 CC microorganisms, especially other *Sphingomonas* strains. The DNA
 CC transferred to other strains includes the spsB gene which encodes a
 CC glycosyl-C55-isoprenylphosphate (glycosyl-1P) transferase. The sphingane
 CC are useful as specialty polymers and as additives in textile
 CC applications, foods, cosmetics, paper, paint, cements, e.g. as viscosity
 CC modifiers, in various other coating applications, and as adhesives and
 CC additives to petroleum products and specialty chemicals.

SQ Sequence 28804 BP; 4974 A; 9804 C; 9230 G; 4796 T; 0 other;

Alignment_scores:

Quality: 215.00 Length: 97
 Ratio: 2.905 Gaps: 0
 Percent Similarity: 76.289 Percent Identity: 45.361

Alignment block:

US-09-484-577A-4 x AAV81474/rev ..

Align seg 1/1 to reverse of: AAV81474 from: 1 to: 28804

```

27 PheleuproAlaAlaLeuGluIleValGluThProProserProthral 43
|||||
18808 TTCTGCGCGCGCGCTCGAGATCATCGAACCTCCGTGTGCCACCGC 18759
43 aArgLeuThAlaAlaLeuLeuAlaLeuPheTYrCysAlaValAla 60
|||||
18758 GCGGCTTACCGCGCGGTGATGTGGCGGCGCTGCGATCACCACCGCT 18709
60 rPalagLeuGluArgIleAspIleValAlaSerAlaSerArglySile 76
|||
18708 GCGTGGCAATCGCGCGCGGTGAGAGTGTGGCGGCGGACGAGCGCGCTATC 18659
77 ValProGlyAspArgValLysLeuValGlnProLeuGluValGlyValVa 93
|||||
18658 GCCCGCATCGCGCGACCAAGATCGTCCGCCGGAAGCGGTATCCT 18609
93 lArgAlaThrHisValArgAspGlyGlnThrValLysAlaGlyGluIle 110
|||||
18608 CCGGCGCATTCGTGTGGCGGAGGCGAGAGTTTCAGAGAGGCGCGAGTGC 18559
110 euIlleGluLeuAspProphealagIlyValAspValAla 123
|||||
18558 TGTATCACGCTCGATCCACCATGTGCGGCGGCGGAGCGCGCGC 18518

```

seq_name: /SISL1/gcdata/hold-geneseq/geneseq-emb1/NA1994.DAT.AA070050

seq_documentation_block:

ID AAQ70050 standard; DNA: 7184 BP.

AC AAQ70050;

DT 13-MAR-1995 (first entry)

DE plx gene of *Pasteurella haemolytica*.

KW *Pasteurella haemolytica*; shipping fever; pasteurellosis; cattle;

KW vaccine; treatment; therapy; prophylaxis; antigen; antibody;

OS immunosay; ss.

OS *Pasteurella haemolytica*.

PH Key Location/Qualifiers

FT CDS 174..3035

```

FT      /*tag= a
FT      /product= PtXA protein

```

PN	US5336491-A.
XX	
PD	09-AUG-1994.
XX	
PF	28-NOV-1986;
XX	
PR	28-NOV-1986;
PR	13-AUG-1987;
PR	18-JUN-1990;
PR	15-JUN-1992;
XX	
PA	(TEXA) UNITV TEXAS SYSTEM.
XX	
PI	Bergert P, Engler M, Highlander S, Weinstock G;
XX	
DR	WPI, 1994-255245/31. P-PSDB; AAK60072.

New purified antigen from *Pasteurella haemolytica* - useful in
 vaccines against shipping fever of cattle, also for raising
 antibodies useful in diagnosis and passive immunisation
 Disclosure; Figure 9, 35pp. English.

Disclosure; Figure 9; 35pp; English.

XX The sequence encodes the PRX protein of *Pasteurella haemolytica*.
CC This protein is a 105 kD antigen which may be used in vaccines to
CC protect cattle against shipping fever. The antigen is also useful
CC for raising antisera which can be used for passive immunisation (for
CC treatment or antiserum prophylaxis) and also to raise antibodies
CC which can be used in immunoassay methods for the detection of
CC which can be used in immunoassay methods for the detection of

xx Sequence 7184 BP; 2366 A; 1267 C; 1485 G; 2066 T; 0 other;
sq

```
alignment_scores:      Length: 103
                        Quality: 190.00
                        Ratio: 2.568
                        Gaps: 0
Percent Similarity: 71.845
Percent Identity: 36.893
```

alignment_block:

```
alignment_block:
```

US-09-484-57/A-4 X AAQ/0050 ..

Align seg 1/1 to: AAQ70050 from: 1 to: /104

[illegible]

```

5647      GCAGATATC 5655
      ::|||:::

```

seq_name: /SIDSL/gcgda/ncbi/genbank/...

seq_documentation_block:
ID_AA006074 standard; DNA; 7183 BP.

AAQ060742

DT 24-JAN-1991 (first entry)

Sequence encoding the look-ahead

VACCINE, SNAPEY -
KM
XX

XX
XX
XX

Qualifiers

FT	1/4..3032
CDS	/*taq= a

XX
XX
DN
HSA957739-A.

XX
PD
18-SEP-1990.

13-AUG-1987; 87US-0085430

PR 13-AUG-1987; 8/05-0005450

PA (TEXA) ONLY OF 1-11-11
XX

XX
F1
2073-1

DR P-PSDB; AAR07167.

PT Vaccine against *Haemophilus influenzae* type b

XX Disclosure; Fig 9a-h; 35pp; English.
ps

Gene product is antigenic to *P. haemolyticus*

CC antigen may be used in passive immunization.

Sequence / 183 BF; 2004 47 2004

```
alignment-scores:      length: 103
      quality:          Gaps: 0
      Ratio: 2.493
Percent Similarity: 70.874 Percent Identity: 35.922
```

alignment_block: AA006074

US-09-484-5/1A-4 X MAG0000/

Align seg 1/1 to: AAQ06074 from: 1 to: 100

```

20  LysThrArgAspGluLeuAlaPheLeuProAlaAlaLeuIleValG1 36
   :: :: :: |||||  |||||  |||||  |||||  |||||  ||
5347  AGAAAAAAGACGAAAGTGAATTTTACCGGCATTTAAGACTGATCA 5396
   :: :: :: |||||  |||||  |||||  |||||  |||||  ||

36  utRProProSerProThrIaArgLeuThrAlaAlaLeuAlaAlaL 53
   |||||  |||  |||||  |||  |||||  ::
5397  AACCCGGTTCTTAAAAAACCCACGCTGAATTCCTATTTCATTGCTAT 5446
   |||||  |||  |||||  |||||  |||||  |||||

53  eupHeTYrCYsAlaValAlaIaTrpAlaGlyLeuGlyArgIleAspIleVal 69
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
5447  TTTTAGTGTGGCATTGTGCTTGCACATGTGAAGAAAGTGAATATGTC 5496
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::

70  AlaSerAlaSerArgLysIleValProGlyAspArgValLysIleValG1 86
   |||||  |||||  |||||  |||||  |||||  |||||

```

```

5497 GCGACTGCTCCCGGTAAATTAATTTAGTGGCAGACGTAAGAAATTTAA 5546
86 nProleugluValglyValAlaValAlaThrHisValArgAspGlyGlnT 103
5547 ACCGATTGAAACGCCATTGTACAGAAATTTTCCTTAAGATGGGCGAGT 5596
103 hrVallysalAGlyGluIleuIleuIleuAspProPheAlaGlyGly 119
5597 TTGTGGAAAGGCGCAATTAATTAAGTCAAGCTTAACGCAATGGCTTGTAT 5646
120 ValAspVal 122
5647 GCAGATATC 5655

seq_name: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT:AA73218
seq_documentation_block:
ID AA73218 standard; DNA; 5120 BP.
XX
XX AA73218;
XX
XX 16-FEB-1998 (first entry)
XX
XX DE ApxIBD gene.
XX
XX KW RTX toxin; apxICA gene; apxIBD gene; apxIAB/C gene; apxIIBCD gene;
XX repeat in toxins toxin; cell-associated RTX toxin; vaccine production;
XX therapy; A. pleuropneumoniae infection; swine pleuropneumonia; ss.
XX OS Actinobacillus pleuropneumoniae.
XX
XX FH Key location/Qualifiers
XX FT -35_signal 524..530
XX FT -10_signal /*tag= a 548..553
XX FT terminator /*tag= b 572..577
XX FT RBS /*tag= c 582..588
XX FT CDS /*tag= d 592..2715
XX FT /*tag= e
XX FT /*product= ApxIB
XX FT RBS 2713..2720
XX FT /*tag= f
XX FT CDS 2724..4160
XX FT /*tag= g
XX FT /*product= ApxID
XX
XX CA2170839-A.
XX
XX PD 02-SEP-1996.
XX
XX PF 01-MAR-1996; 96CA-2170839.
XX
XX PR 01-MAR-1995; 95US-0396244.
XX
XX PA (UYGV-) UNITV GUELPH.
XX
XX PI Macinnes J, Mallard B, Ricciatti P, Rosendal S;
XX
XX DR WPI; 1997-245536/23.
XX
XX DR P-PSDB; AAW22153, AAW22154.
XX
XX PT Preparations of microorganisms producing cell-associated RTX toxins
XX pleuro-pneumonia
XX
XX PS Disclosure: Pages 97-100; 151pp; English.
XX
XX This sequence represents the A. pleuropneumoniae apxIBD gene. This
XX sequence, and the apxICA, apxIAB/C, and apxIIBCD genes (see
XX AA73217, AA73219 and AA73220) can be expressed by microorganisms used

```

```

CC in the preparations of the invention. The preparations are bacterial
CC microorganisms comprising one or more isolated and purified strains of a
CC microorganism that produces one or more RTX (repeat in toxins) toxins,
CC where the strains have at least one cell-associated RTX toxin. The
CC preparations are used for production of vaccines for the prophylaxis and
CC treatment of infectious diseases caused by microorganisms that produce
CC RTX toxins, where the strains have been attenuated or inactivated. The
CC vaccines are preferably against Actinobacillus pleuropneumoniae
CC infection (swine pleuropneumonia). It has been found that A.
CC pleuropneumonia produces significant quantities of cell-associated RTX
CC toxins when cultured under certain conditions, and that the whole-cell
CC protein profiles obtained from the cultures corresponds to the whole-cell
CC protein profiles obtained from cells recovered at necropsy from the
CC pleural fluid of infected swine. Vaccination with a bacterin from the
CC from heat-inactivated cultures having significant quantities of
CC cell-associated RTX toxins give significant protection of swine against
CC challenge with homologous strains.
XX
XX Sequence 5120 BP; 1580 A; 880 C; 1139 G; 1521 T; 0 other;
XX

```

```

alignment_scores:
Quality: 169.00 Length: 102
Ratio: 2.315 Gaps: 0
Percent Similarity: 71.569 Percent Identity: 35.294

```

```

alignment_block:
US-09-484-577A-4 x AA73218

```

```

Align seg 1/1 to: AA73218 from: 1 to: 5120

```

```

20 LysThrArgAspGluLeuAlaPheLeuProAlaAlaLeuGluIleValGln 36
2823 CGAGAAAGAGATGAAATGATTTTACTGCACTTAAGCTGATTTGA 2872
36 thrProPseProThrAlaArgLeuThrAlaAlaLeuAlaVal 53
2873 AACACCGGTGTAAAAAACCGAGTGTGCTTAATTAATTTGCGT 2922
53 eupheryCysAlaValAlaThrPalaglyLeuGlyArgGlyLeuVal 69
2923 TCCATTTTGGCATTTAGTTTCAATTTGACGACGACGACGACGACG 2972
70 AlaSerAlaSerArgGlyIleValProGlyAspArgValIleuValGln 86
2973 GCGACCGCAACGGTAAATTAATTAATTAATTAATTAATTAATTA 3022
86 nProleugluValglyValAlaValAlaThrHisValArgAspGlyGlnT 103
3023 GCCGATTGAAACGCCCTTGGTTAAAGAAATCTTGTGCAAGCGCAAT 3072
103 hrVallysalAGlyGluIleuIleuIleuAspProPheAlaGlyGly 119
3073 TTGTGGAAAGATCAAGTGTGTTACCTTGACCGCATTTGGAGCGCAT 3122
120 ValAsp 121
3123 GCGGAT 3128

seq_name: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:AA288585
seq_documentation_block:
ID AA288585 standard; DNA; 5120 BP.
XX
XX AC AA288585;
XX
XX DT 05-MAY-2000 (first entry)
XX
XX DE A. pleuropneumoniae DNA encoding clyTB and clyTD.
XX
XX KW RTX toxin; vaccine; Repeats in toxins; immunostimulatory; antimicrobial;
XX anti-inflammatory; antidiarrhetic; antidiarrhetic; treatment; pneumonia;
XX pleuropneumonia; septicemia; nephritis; arthritis; endocarditis;

```


[illegible]

PI Macinnes J, Mallard B, Ricciatti P, Rosendal S;
 XX
 XX WPI: 1997-245536/23.
 DR P-PSDB: AAMW22158, AAMW22159, AAMW22160, AAMW22161.
 DR
 XX Preparations of microorganisms producing cell-associated RTX toxins
 PT - especially for production of vaccines against swine
 PT pleuro-pneumonia
 XX
 PS Disclosure; Pages 114-118; 151pp; English.
 XX

CC virus sequence represents the A. pleuropneumoniae apxIIABCD gene. This
 CC sequence, and the apxICA, apxIIB, and apxIIB/C genes (see
 CC AAI73317-733119) can be expressed by microorganisms used in the
 CC preparations of the invention. The preparations are bacterial
 CC preparations comprising one or more isolated and purified strains of a
 CC microorganism that produces one or more RTX (repeat in toxins) toxins,
 CC where the strains have at least one cell-associated RTX toxin. The
 CC preparations are used for production of vaccines for the prophylaxis and
 CC treatment of infectious diseases caused by microorganisms that produce
 CC RTX toxins, where the strains have been attenuated or inactivated. The
 CC vaccines are preferably against *Acetivibacillus pleuropneumoniae*
 CC infection (swine pleuropneumonia). It has been found that A.
 CC pleuropneumonia produces significant quantities of cell-associated RTX
 CC toxins when cultured under certain conditions, and that the whole-cell
 CC protein composition of the cultures corresponds to that the whole-cell
 CC protein profiles obtained from cells recovered at necropsy from the
 CC pleural fluid of infected swine. Vaccination with a bacterin prepared
 CC from heat-inactivated cultures having significant quantities of
 CC cell-associated RTX toxins give significant protection of swine against
 CC challenge with homologous strains.

CC Sequence 7721 BP; 2593 A; 1146 C; 1520 G; 2462 T; 0 other;
 CC 50

```
alignment_scores:
  Quality: 168.00      Length: 1022
  Ratio: 2.333         Gaps: 0
Percent Similarity: 70.588      Percent Identity: 35.294
```

segment_block:
09-484-577A-4 x AAT73220 . .

Align seg 1/1 to: AAT73220 from: 1 to: 7721

```

12  ArgGluValValAlaValaIgiLyLysThrArgAspGluLeuAlaPhe 28
1615 CGCAAAACATTAGATGACCACGAAACAAAAAGATGAACGAATTTT 6214
28  uProAlaAlaLeuGluIleValaIuThrProPseProThrAlaArgL 45
6215 GCCCGCGCATTTAGATTAATTGACACCGCTATTTCAAAAAGCCACGCG 6264
45  euThrAlaAlaLeuLeuAlaAlaLeuPheTyCysAlaValaIaTrpIa 61
6265 TGATCGCTTTTGGATTAACGCTATTTCTATTTTACATTTGTAAATTTC 6314
62  GlyLeuGlyArgIleAspIleValaAlaSerAlaSerArgLysIleValaPr 78
6315 ATTATTGTGTAAGATGAAGATTGTGCTACTGCTACAGGTAAGTGGTAT 6364
78  ogLysAspArgValLysLeuValaIleProLeuGluValaGlyValaIaArg 95
6365 TAGGGGACATGTAAGAAGAATAAAGCCTTTGAGAACTGTTAGTAAGAAG 6414
95  IaThrHisValaIaArgAspGlyGlnThrValaLysAlaGlyGluIleuLeu 111

```

Tue Jul 16 05:46:31 2002

us-09-484-577a-4.rng

Page 13



OM of: US-09-484-577A-4 to: Issued_Patents_NA:* out-format : pfs
Date: Jul 16, 2002 2:11 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=framed_p2n.model -DEV=x1h
-CGN2_1/USPTO.spool/US09484577/rn1at.15072002.153753.8157/app_query.fasta_1.184
-DB=Issued_Patents_NA -OPMT=fastap -SUFFIX=rni -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOPCL=0.000 -LOPEXT=0.000
-OGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -DELEX=7.000 -YGAPO=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEX=7.000 -START=1 -MATRIX=blotsum62
-TRANS=human40.cdi -LIST=45 -DOCCALIGN=200 -THR_SCORE=Pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTEMT=pfs
-NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09484577_@CGN1_1_45 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPEX -WAIT -THREADS=1

Search information block:

Query: US-09-484-577A-4
Query length: 124
Database: Issued_Patents_NA.*
Database sequences: 383533
Database length: 122816752
Search time (sec): 42.840000

score_list:
Sequence Strd Orig ZScore EScore Len Documentation
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-592-874-1- 215.00 411.21 3.2e-15 28804
/cgn2_6/pdata/2/ina/6A.COMB.seq:US-09-096-842-2- 215.00 411.21 3.2e-15 28804
/cgn2_6/pdata/2/ina/6A.COMB.seq:US-09-096-867-2- 215.00 411.21 3.2e-15 28804
/cgn2_6/pdata/2/ina/6A.COMB.seq:US-08-772-270A-6 + 169.00 329.26 1.2e-10 5120
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-488-706-1 + 169.00 329.26 1.2e-10 5120
/cgn2_6/pdata/2/ina/6A.COMB.seq:US-08-772-270A-14 + 169.00 329.26 1.2e-10 5120
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-620-605D-1 + 154.00 300.34 4.9e-09 3576
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-620-605D-1 + 99.00 177.78 0.0326 4465
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-09-491-772-5 + 99.00 177.58 0.0334 4547
/cgn2_6/pdata/2/ina/6B.COMB.seq:US-09-491-772-5 + 90.00 171.35 0.0743 1302
/cgn2_6/pdata/2/ina/6B.COMB.seq:US-09-491-772-1 + 90.00 151.10 0.9972 8580
/cgn2_6/pdata/2/ina/6B.COMB.seq:US-09-614-770A-1 + 77.00 103.92 4.23.25 49272
/cgn2_6/pdata/2/ina/5A.COMB.seq:US-08-614-770A-17 + 75.50 137.84 5.46 1543
/cgn2_6/pdata/2/ina/5A.COMB.seq:US-09-614-770A-17 + 75.50 137.84 5.46 1543
/cgn2_6/pdata/2/ina/5A.COMB.seq:US-08-188-582-15 + 74.50 126.55 23.23 3603
/cgn2_6/pdata/2/ina/5A.COMB.seq:US-08-646-715-15 + 74.50 126.55 23.23 3603
/cgn2_6/pdata/2/ina/6B.COMB.seq:US-09-381-862-5 + 73.00 116.63 82.93 6688
/cgn2_6/pdata/2/ina/6B.COMB.seq:US-09-381-862-5 + 73.00 116.63 82.93 6688
/cgn2_6/pdata/2/ina/6B.COMB.seq:US-09-103-840A-2 - 72.00 103.05 473.04 23673
/cgn2_6/pdata/2/ina/5A.COMB.seq:US-07-640-476-4 - 70.50 129.78 15.35 1185
/cgn2_6/pdata/2/ina/5A.COMB.seq:US-07-640-476-4 - 70.50 129.78 15.35 1185
/cgn2_6/pdata/2/ina/6A.COMB.seq:US-08-458-568A-11 + 70.50 126.76 22.62 1566
/cgn2_6/pdata/2/ina/5A.COMB.seq:US-09-103-840A-1 - 69.50 104.89 373.90 12001
/cgn2_6/pdata/2/ina/6B.COMB.seq:US-09-103-840A-1 - 69.50 104.89 373.90 12001
/cgn2_6/pdata/2/ina/6B.COMB.seq:US-08-387-942C-1 - 69.00 101.10 607.96 12588
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-804-227C-7 - 69.00 87.57 3.4e+03 44377
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-804-227C-7 - 69.00 87.57 3.4e+03 44377
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-103-840A-1 + 69.00 87.57 3.4e+03 44377
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-07-661-610C-7 + 68.50 113.43 124.98 3606
/cgn2_6/pdata/2/ina/5A.COMB.seq:US-09-199-737-3 + 68.00 122.37 39.71 1417
/cgn2_6/pdata/2/ina/6B.COMB.seq:US-08-922-865-1 - 67.50 117.84 70.97 1951
/cgn2_6/pdata/2/ina/6A.COMB.seq:US-08-922-865-1 - 67.50 117.84 70.97 1951
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-809-763-5 - 67.50 117.77 71.68 1965
/cgn2_6/pdata/2/ina/6A.COMB.seq:US-08-809-763-5 - 67.50 117.77 71.68 1965
/cgn2_6/pdata/2/ina/6A.COMB.seq:US-08-286-870A-7 - 67.00 116.76 81.60 2159
/cgn2_6/pdata/2/ina/5A.COMB.seq:US-08-286-870A-7 - 67.00 116.76 81.60 2159
/cgn2_6/pdata/2/ina/5A.COMB.seq:US-08-090-075-3 - 66.00 125.53 26.49 703
/cgn2_6/pdata/2/ina/6B.COMB.seq:US-09-219-337-1 + 66.00 106.85 290.70 4002
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-231-193A-53 + 66.00 106.85 290.70 4002
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-486-273A-53 + 66.00 106.85 290.70 4002
/cgn2_6/pdata/2/ina/6A.COMB.seq:US-08-480-474-53 + 66.00 106.85 290.70 4002

/cgn2_6/pdata/2/ina/6A.COMB.seq:US-08-940-086A-53 + 66.00 106.85 290.70 4002
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-940-035A-53 + 66.00 106.85 290.70 4002
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-231-193A-49 + 66.00 106.81 292.20 4017
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-486-273A-49 + 66.00 106.81 292.20 4017
/cgn2_6/pdata/2/ina/6A.COMB.seq:US-08-480-474-49 + 66.00 106.81 292.20 4017

seq_name: /cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-592-874-1

seq_documentation_block:

Sequence 1, Application US/08592874
Patent No. 5854034
GENERAL INFORMATION:
APPLICANT: POLLOCK, THOMAS J.
APPLICANT: YAMAZAKI, MOTOHIDE
APPLICANT: THORNE, LINDA
APPLICANT: MIKOLAJCZAK, MARCIA
APPLICANT: ARMENTROUT, RICHARD W.
TITLE OF INVENTION: DNA SEGMENTS AND METHODS FOR INCREASING
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: JULES E. GOLDBERG
STREET: 261 MADISON AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016-2391
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,874
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/377,440
FILING DATE: 24-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: GOLDBERG, JULES E.
REGISTRATION NUMBER: 24,408
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-986-4090
TELEFAX: 212-818-9479
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 28804 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FRAGMENT TYPE: N-terminal
US-08-592-874-1

alignment_scores:

Quality: 215.00 Length: 97
Ratio: 2.905 Gaps: 0
Percent Similarity: 76.289 Percent Identity: 45.361

alignment_block:

US-09-484-577A-4 x US-08-592-874-1/rev ..

Align seg 1/1 to reverse of: US-08-592-874-1 from: 1 to: 28804

27 PheLeuProAlaAlaLeuGluIleValGluThrProSerProThrAl 43

18808 TTCCTGCGCGCGCGCGTCGATCATCGACGTCGCGTCCGCCACCGC 18759

43 argLeuThrAlaAlaLeuAlaAlaAlaLeuPheTyrCysAlaValAla 60

18758 GCGGCTCACCGCGCGGTGATGTGGCGGCGGCGCATCATCACCGCCT 18709

```

seq_name: /cgn2_6/ptodata/2/1na/6A_COMB.seq:US-09-096-867-2
seq documentation block:
; Sequence 2, Application US/09096867
; Patent No. 6030817
; GENERAL INFORMATION:
; APPLICANT: Pollock, Thomas J
; APPLICANT: Mikolajczak, Marcia
; APPLICANT: Yamazaki, Motohide
; APPLICANT: Thorne, Linda
; APPLICANT: Amentrout, Richard W
; TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteriaria
; FILE REFERENCE: seq list for appl filed from Xanthomonas Campestris
; CURRENT APPLICATION NUMBER: US/09/096,867
; CURRENT FILING DATE: 1998-06-11
; EARLIER APPLICATION NUMBER: 60/049,428
; EARLIER FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 28804
; TYPE: DNA
; ORGANISM: Sphingomonas sp. 588
; US-09-096-867-2

```

[illegible]

```
APPLICANT: Ricciatti, Paul
APPLICANT: Mallard, Bonnie
APPLICANT: Rosendal, Soren
TITLE OF INVENTION: NOVEL BACTERIAL PREPARATIONS, METHOD FOR
TITLE OF INVENTION: PRODUCING SAME, AND THEIR USE AS VACCINES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Bereskin & Parr
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/772,270A
FILING DATE: December 23, 1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Gravelle, Michelle
REGISTRATION NUMBER: 40,261
REFERENCE/DOCKET NUMBER: 6580-81
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 5120 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Actinobacillus pleuropneumoniae
US-08-772-270A-6

alignment_scores:
Quality: 169.00 Length: 102
Ratio: 2.315 Gaps: 0
Percent Similarity: 71.569 Percent Identity: 35.294

alignment_block:
US-09-484-577A-4 x US-08-772-270A-6 ..
Align seg 1/1 to: US-08-772-270A-6 from: 1 to: 5120

20 LysThrArgAspGluLeuAlaPheLeuProAlaAlaLeuGluIleValG1 36
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2823 CGAGAAAGATGAAATGAATTTTACCTGCACACTTAGAGCTAATTGA 2872
36 uThrProSerProThrAlaArgLeuThrAlaAlaLeuLeuAlaAlaL 53
||||| ||| ||||| ||| |||||
2873 AACACCGGTGCACAAAACCGAGATGATGCTTATTATTAATGCGT 2922
53 eupheTyTCysAlaValAlaTrpAlaGlyLeuGlyArgIleAspIleVal 69
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2923 TCCTATTTTGGCATTAGTTATTTCATGTGTCAGTCACGTAGAAATTTG 2972
70 AlaSerAlaSerArgIleValAlaProGlyAspArgValIleLeuValG1 86
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2973 GCGACCGCACCGGTAAATTAATGCGTTAGCGACCGTAGCAAGAAATTAA 3022
86 nProLeuGluValGlyValAlaArgAlaThrHisValArgAspGlyGlnT 103
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3023 GCGGATGAAACGCGCTGTGAAGAAATCTTTGCAAGACGAGCAAT 3072
103 hValIysAlaGlyIleLeuIleGluLeuAspProPheAlaGlyG1 119
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
3073 TTGTTGAGAAAGATCACTGCTGTACACTTGACCGCATTTGGAGCCGAT 3122
120 ValAsp 121
::|||
3123 GCGGAT 3128

seq_name: /cgn2_6/ptodata/2/lna/5B.COMB.seq:US-08-488-706-1
seq_documentation_block:
; Sequence 1, Application US/08488706
; Patent No. 5994525
; GENERAL INFORMATION:
; APPLICANT: Kamp, Elbarie M.
; APPLICANT: Smits, Marinus A.
; TITLE OF INVENTION: Recombinant Vaccine For Prevention and/or
; TITLE OF INVENTION: Treatment of Pleuropneumonia Infections
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Abelman, Frayne & Schwab
; STREET: 708 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM PC or compatible
; OPERATING SYSTEM: DOS 3.31
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,706
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/722,971
; FILING DATE: 28-JUNE-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Cinnamon, Jay S.
; REGISTRATION NUMBER: 24,156
; REFERENCE/DOCKET NUMBER: 201,875
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-949-9022
; TELEFAX: 212-949-9190
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8370
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-488-706-1

alignment_scores:
Quality: 169.00 Length: 102
Ratio: 2.315 Gaps: 0
Percent Similarity: 71.569 Percent Identity: 35.294

alignment_block:
US-09-484-577A-4 x US-08-488-706-1 ..
Align seg 1/1 to: US-08-488-706-1 from: 1 to: 8370

20 LysThrArgAspGluLeuAlaPheLeuProAlaAlaLeuGluIleValG1 36
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
6073 CGAGAAAGATGAAATGAATTTTACCTGCACACTTAGAGCTAATTGA 6122
36 uThrProSerProThrAlaArgLeuThrAlaAlaLeuLeuAlaAlaL 53
||||| ||| ||||| ||| |||||
6123 AACACCGGTGCACAAAACCGAGATGATGCTTATTATTAATGCGT 6172
53 eupheTyTCysAlaValAlaTrpAlaGlyLeuGlyArgIleAspIleVal 69
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
6173 TCCTATTTTGGCATTAGTTATTTCATGTGTCAGTCACGTAGAAATTTG 6222
```

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-772-270A-14

GENERAL INFORMATION:
APPLICANT: Macinnes, Janet
APPLICANT: Ricciatti, Paul
APPLICANT: Mallard, Bonnie
APPLICANT: Rosendal, Soren
TITLE OF INVENTION: NOVEL BACTERIAL PREPARATIONS, METHOD FOR
TITLE OF INVENTION: PRODUCING SAME, AND THEIR USE AS VACCINES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:

ADDRESSEE: Bereskin & Parr
 STREET: 40 King Street West
 CITY: Toronto
 STATE: Ontario
 COUNTRY: Canada
 ZIP: M5H 3Y2
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.33
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/772,270A
 FILING DATE: December 23, 1996
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: Gravelle, Michelle
 REGISTRATION NUMBER: 40,261
 REFERENCE/DOCKET NUMBER: 6580-81
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (416) 364-7311
 TELEFAX: (416) 361-1398
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 7721 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 ORIGINAL SOURCE:
 ORGANISM: Actinobacillus pleuropneumoniae
 OS-08-772-270A-14

Quality:	168.00	Length:	102
Ratio:	2.333	Gaps:	0
Percent Similarity:	70.588	Percent Identity:	35.294

alignment_block:
US-09-484-577A-4 x US-08-772-270A-14

```

12 ASGGILVALVALAVALGLYGLYSTRHRRGASPGILLEUALALEPHE 28
   |||:::  :::  :::  :::  |||:::  |||:::  |||:::  |||:::
6165 CGGAAACAATTGATGATACCCACCAAGCAAAAGATGMAACGATTTT 6214
   |||:::  :::  :::  :::  |||:::  |||:::  |||:::  |||:::
28 UPROMAALALEUGLILEVALGLUTHRPROSESPROTHRILARGL 45
   |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
6215 GCCTCGGCATTTAGAGTTAATTGACACACCATATTTCAAAAAAGCCAGCC 6264
   |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
45 EUTHRALALEULEUALALEUPHETRYCYSAVALALETRPALA 61
   |||:::  |||:::  :::  :::  :::  :::  :::  :::
6265 TGAATGCTATTGTGATATGCAATTCATTTTATTTAGCATATGTAATTTC 6314
   |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
62 GLYLEUGLYAIRILASPLILEVALILASERILASERATGLYSILEALPR 78
   |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
6315 ATTATTAGTAAGTAAAGTAAATTTCTCTAGTGTGCTAAGGTGATTT 6364
   |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
78 OGILSPARGVALYSLEULVALGINPROLEUGLIVAGLYVALVALARGA 95
   :::  :::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
6365 TNGTGACATAGTAAAGAAATTAACCCATTAAGAGATGCTTACTATAAAG 6414
   |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
95 IATHHISVALARGASPGLYGLNTHRVALLYSALAGLYGLULLEULE 111
   |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
6415 ACAATTGTTTAAAGATGACACAAATTTGTTGAAAAAGCACATATTTATTA 6464
   |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
112 GLULEU 113
   |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
6465 AATCTC 6470

```

seq_name: /cgn2_6/ptodata/2/lna/6B_COMB.seq:US-08-976-259-79

```
seq_documentation_block;
```

; Sequence 79, Application US/08976255
; Patent No. 6316609

```

; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.

```

APPLICANT: Choi, Gill H.
APPLICANT: Welch, Rodney A.

TITLE OF INVENTION: Nucleotide Sequence of Escherichia coli
Patent No. 6316609

NUMBER OF SEQUENCES: 142
CORRESPONDENCE ADDRESS:

ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Ave, N.W., Suite 600

CITY: Washington
STATE: DC
COUNTRY: USA

COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
CONNECTION: via RS-232C

```

;
COMPUTER:  HP Vectra 486/33
;
OPERATING SYSTEM:  MSDOS version 6.2
;
SOFTWARE:  ACCT 1994

```

```

; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: HSC/08/0756 350

```

APPLICATION NUMBER: 05/08/910,235
FILING DATE: Herewith
OF APPLICATION, 536

CLASSIFICATION: 330
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/031 636 AND US 60/061 051

ATTORNEY/AGENT INFORMATION
NAME: STAFFORD, ERIC R

NAME: STEELE, ELLIC A.
REGISTRATION NUMBER: 36,688
REFERENCE / DOCKET NUMBER: 1488 0740003 / ETC / 0000

REFERENCE/DOCKET NUMBER: 14
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-3600

TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR CDO TD NO. 70

SEQUENCE CHARACTERISTICS:

LENGTH: 3570 base pairs
TYPE: nucleic acid
COMMENTARY: 4 sub

```

;
;
; STRANDEDNESS: double
; TOPOLOGY: linear
;

```



```

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)-205-8000
TELEFAX: (703)-205-8050
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4547 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
ORIGINAL SOURCE:
SPRAIN: Serratia marcescens Str1
US-09-005-232A-1

```

```

alignment_scores:
Quality: 99.00      Length: 57
Ratio: 2.538        Gaps: 2
Percent Similarity: 68.421  Percent Identity: 43.860

```

```

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-491-772-5

```

```

Align seg 1/1 to: US-09-005-232A-1 from: 1 to: 4547
58 ValAlaTTPAlaGlyLeuGlyArgGlyLeuAspIle...ValAlaSerAlaSe 73
:::|||||:::|||||:::|||||:::|||||:::
1936 TTACCTGGGGGGGCTTGGCGCTTGATAAAGGGGTGGCGCGCGG 1985
73 rArgLysIleValProGlyAspArgValLysLeuValGlnProLeuGluV 90
||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1986 ATCGTAAACCGCTTCCGGCAACGCC...AAACGGTCAGAGCCCGGCCCA 2032
90 alGlyValValArgAlaThrHisValArgAspGlyGlnThrValLysAla 106
|||||:::|||||:::|||||:::|||||:::
2033 GCGGCATCATTAAGATATTCGGCTCAGAGATGGCACAAGTGAAGACC 2082
107 GlyLysIleLeuIleGluLeu 113
|||||:::|||||:::|||||:::
2083 GGTGAGTCTGCTGTCAGCTC 2103

```

```

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-491-772-5
seq_documentation_block:
; Sequence 5, Application US/09491772
; Patent No. 6329172
; GENERAL INFORMATION:
; APPLICANT: Rhee, Joon Shick
; APPLICANT: Pan, Jae Gu
; APPLICANT: Ahn, Jung Hoon
; TITLE OF INVENTION: ABC Transporter Gene Cluster in Pseudomonas fluorescens for
; FILE REFERENCE: Enhanced Lipase Secretion
; CURRENT APPLICATION NUMBER: US/09/491,772
; PRIOR FILING DATE: 2000-01-26
; PRIOR APPLICATION NUMBER: KR 10-1999-0008512
; NUMBER OF SEQ ID NOS: 1999-03-13
; SOFTWARE: KOPATIN 1.5
; SEQ ID NO 5
; LENGTH: 1302
; TYPE: DNA
; ORGANISM: Pseudomonas fluorescens
US-09-491-772-5

```

```

alignment_scores:
Quality: 90.00      Length: 63
Ratio: 2.250        Gaps: 0
Percent Similarity: 63.492  Percent Identity: 36.508

```

```

alignment_block:

```

```

US-09-484-577A-4 x US-09-491-772-5 ..
Align seg 1/1 to: US-09-491-772-5 from: 1 to: 1302

```

```

60 TTPAlaGlyLeuGlyArgGlyLeuValAlaSerAlaSerArgLysI 76
|||||:::|||||:::|||||:::|||||:::
94 TGGCCAGCGCTGGCGCCGTCGACCAAGGCAATTCGGTGAAGCAACCGT 143
76 eValProGlyAspArgValLysLeuValGlnProLeuGluValGlyValV 93
||| ::| ::| ||| ::| ::| ::| ::| ::| ::| ::| ::|
144 CGTGGCTCGGCAAGCGCAAGCGGTGCAACCTTCAGCCCGGCGGTGG 193
93 alArgAlaThrHisValArgAspGlyGlnThrValLysAlaGlyLysIle 109
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
194 TCAGCCGATTCGTGGTCGCGCAGGCGCAACGTTAAACAAGCCAGCGCG 243
244 CTGTTTCGCTGACCAAGCCAGCAAGCGCTGATGTG 282

```

```

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-614-770A-1

```

```

seq_documentation_block:
; Sequence 1, Application US/09491772
; Patent No. 6329172
; GENERAL INFORMATION:
; APPLICANT: Rhee, Joon Shick
; APPLICANT: Pan, Jae Gu
; APPLICANT: Ahn, Jung Hoon
; TITLE OF INVENTION: ABC Transporter Gene Cluster in Pseudomonas fluorescens fo
; FILE REFERENCE: Enhanced Lipase Secretion
; CURRENT APPLICATION NUMBER: US/09/491,772
; PRIOR FILING DATE: 2000-01-26
; PRIOR APPLICATION NUMBER: KR 10-1999-0008512
; NUMBER OF SEQ ID NOS: 1999-03-13
; SOFTWARE: KOPATIN 1.5
; SEQ ID NO 1
; LENGTH: 8580
; TYPE: DNA
; ORGANISM: Pseudomonas fluorescens
US-09-491-772-1

```

```

alignment_scores:
Quality: 90.00      Length: 63
Ratio: 2.250        Gaps: 0
Percent Similarity: 63.492  Percent Identity: 36.508

```

```

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-614-770A-1

```

```

Align seg 1/1 to: US-09-491-772-1 from: 1 to: 8580
60 TTPAlaGlyLeuGlyArgGlyLeuValAlaSerAlaSerArgLysI 76
|||||:::|||||:::|||||:::|||||:::
4085 TGGCCAGCGCTGGCGCCGTCGACCAAGGCAATTCGGTGAAGCAACCGT 4134
76 eValProGlyAspArgValLysLeuValGlnProLeuGluValGlyValV 93
||| ::| ::| ||| ::| ::| ::| ::| ::| ::| ::| ::|
4135 CGTGGCTCGGCAAGCGCAAGCGGTGCAACCTTCAGCCCGGCGGTGG 4184
93 alArgAlaThrHisValArgAspGlyGlnThrValLysAlaGlyLysIle 109
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4185 TCAGCCGATTCGTGGTCGCGCAGGCGCAACGTTAAACAAGCCAGCGCG 4234
110 LeuIleGluLeuAspProPheAlaGlyValAlaAspVal 122
|||||:::|||||:::|||||:::|||||:::
4235 CTGTTTCGCTGACCAAGCCAGCAAGCGCTGATGTG 4273

```

```

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-614-770A-1

```

```

Alignment_scores:
      Quality: 77.00      Length: 108
      Ratio: 1.453      Gaps: 4
      Percent Similarity: 49.074      Percent Identity: 29.630

Alignment_block:
US-09-484-577A-4 x US-08-614-770A-1 ..

Align seg 1/1 to: US-08-614-770A-1 from: 1 to: 49272

14 ValValaIaValaIGlyGIySthrRArgspjGuleAlaPheLeuProAl 30
||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
41841 GtGATCCCGGTTGAAACACACGATCCGGCGCGGATTCGAATTGTCCAC 41890
30 aAlaLeu.....GluIleValaGluT 37
|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
41891 GsCCCTTGCTCAGGTTCTCTGAGGCGCACGTCGGCGTGGAGGTACACCGCA 41940
37 hrProPterProthrRAlaArgLeuThrAlaIleLeuAlaLeuAlaLeu 53
|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
41941 CCACGCCAGACGGCGGCGAGGCGGATGCGTGGCGGTGGCGATGTCCTTC 41990
54 PheTYrYSAlaValaAlaTrp.....AlaGlyLeuGlyArgIleAspI 68
|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
41991 TTCATGTCGCGTACTTCTTGTTGGTGGCTGTCAAGTACGATGATCGAAGT 42040

```

seq_name: /cgn2_6

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-188-582-15

```
seq_documentation_block:
; Sequence 15, Application US/08188582
; Patent No. 5534410
GENERAL INFORMATION:
APPLICANT: Tjian, Robert
APPLICANT: Comal, Lucio
APPLICANT: Dynlacht, Brian D.
APPLICANT: Hoey, Timothy
APPLICANT: Ruppert, Siegfried
APPLICANT: Tanese, Naoko
APPLICANT: Wang, Edith
APPLICANT: Weinzierl, Robert O.J.
TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOEBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/188,582
FILING DATE: 28-JAN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 3603 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2214
US-08-188-582-15

alignment_scores:
Quality: 74.50 Length: 117
Ratio: 1.202 Gaps: 4
Percent Similarity: 52.991 Percent Identity: 27.350

alignment_block:
US-09-484-577a-4 x US-08-188-582-15 ..
Align seg 1/1 to: US-08-188-582-15 from: 1 to: 3603

4 LeuserProArqmetlysserAlaArgGluValAlaValGlyGly 20
111 111 ..... 111111
913 CTGTGTCCTTCTCTGAG.....AGAGCTTACCGCGCTTGAGACAGCT 956
20 sThArGAspGluLeuAlaPheLeuProAlaAlaLeuGluLeuValGlu 37
111 111111 ..... 111111
957 GAACCGCGACTCGCGCGCTTCAATCG.....CAGAGCCAGCAGC 997
37 hrProSerProThrAlaArgLeuThrAlaAlaLeuAlaAlaLeu 53
111111 111111111111 111111111111
```

```
998 AGCGCCACCGCCACCTCGACGCCACCACTGCGCTACAGCGCGCTGTG 1047
54 PheTyrCysAlaValAlaTrpAlaGlyLeuGlyArgGlyIleAspIleValAl 70
1048 CTGAGTAGCTCGGTCCAGCGCAGCGCGGAGAGAGCGGCCACCGCTGAC 1097
70 aserAlaSerArgLysIleValProGlyAspArgValLysLeuValGlnP 87
1098 CACTGCCCTCCAGCCCCCTGTG.....CTCAGCTCAGCAGC 1135
87 roLeuGluValGlyValAlaArgAlaThrHisValAlaArgAspGlyGlnThr 103
1136 CCACGAGGTCTGCGTC.....GCAAGCAG 1161
104 ValLysAlaGlyLeuLeuLeuGluLeuAspProPheAlaGlyGlyVal 120
1162 GGGCACCCACACCGCTGTGTCATCCAGAGCCTCGACAGCAGGACCCCT 1211
120 1 120
1212 G 1212
```

```
seq_name: /cgn2.6/ptodata/2/1na/5A_COMB.seq:US-08-646-715-15
seq_documentation_block:
; Sequence 15, Application US/08646715
; Patent No. 5637686
GENERAL INFORMATION:
APPLICANT: Tjian, Robert
APPLICANT: Comal, Lucio
APPLICANT: Dynlacht, Brian D.
APPLICANT: Hoey, Timothy
APPLICANT: Ruppert, Siegfried
APPLICANT: Tanese, Naoko
APPLICANT: Wang, Edith
APPLICANT: Weinzierl, Robert O.J.
TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOEBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/646,715
FILING DATE: 09-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/188,582
FILING DATE: 28-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 3603 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
```

FEATURE:
NAME/KEY: CDS
LOCATION: 1..2214
US-08-646-715-15

alignment_scores:

Quality: 74.50 Length: 117
Ratio: 1.202 Caps: 4
Percent Similarity: 52.991 Percent Identity: 27.350

alignment block:

US-09-484-577a-4 x US-08-646-715-15 ..

Align seg 1/1 to: US-08-646-715-15 from: 1 to: 3603

```
4 LeuSerProArgMetIysSerAlaArgGluValAlaAlaValGlyGly 20
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
913 CTTGTGCTTCTTCTGAG.....AGGAGCTTACCCGCTTGAGACAGCT 956

20 sthrArgaspGluLeuAlaPheLeuProAlaAlaLeuGluIleValGluT 37
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
957 GACCCCGGACTCGCGGCTTCATCCAG.....CAGAGCCAGCAGC 997

37 hrProProSerProThrAlaArgLeuThrAlaAlaLeuAlaAlaLeu 53
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
998 AGCCGCCACCGCCACCTCGCAGCGCCAGCAGCTCGCGCTGCGCGTGTG 1047

54 PheTyrCysAlaValAlaIleValAlaGlyLeuGlyArgIleAspIleValAl 70
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1048 CTGAGTACCTCGGTCCAGCGCGCCGCGGAGAGCGCGGCCACCGTGC 1097

70 aSerAlaSerArgIleValProGlyAspArgValIysLeuValGlnP 87
:||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1098 CAGTCCCTCCAGCCCCCTGTG.....CTCAGCCTCAGCGCAGC 1135

87 roLeuGluValGlyValValAlaArgAlaThrHisValAlaArgaspGlyInThr 103
|| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1136 CCACGCAAGTGGCGTC.....GGCAAGCAG 1161

104 ValIysAlaGlyGluIleLeuIleGluLeuAspProPheAlaGlyIysVal 120
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1162 GGGCAACCCACACCGCTGTATCCAGCAGCTCCGAAAGCCAGAGCCCT 1211

120 1 120
1212 G 1212
```

